



Sisältö:

1. *Johdanto*
2. *Yleistetyt
käänteiskuvaukset*
3. *Tihonov-
regularisointi*
4. *Iteratiiviset
menetelmät*
5. *Tehtäviä*

Etusivu



Sivu 1 / 108

Takaisin

Full Screen

Sulje

Poistu

Käänteisongelmat

Jari P. Kaipio

1. elokuuta 2002



Sisältö:

1. *Johdanto*
2. *Yleistetyt käänteiskuvaukset*
3. *Tihonov-regularisointi*
4. *Iteratiiviset menetelmät*
5. *Tehtäviä*

Etusivu



Sivu 2 / 108

Takaisin

Full Screen

Sulje

Poistu

1	Johdanto	3
1.1	Suorat ongelmat ja käänteisongelmat	3
1.2	Esimerkkejä suorista ongelmista	4
1.3	Ratkaisuhilan valinta ja derivointi	12
1.4	Katkaistut hajoitelmat ja “terminen arkeologia”	14
2	Yleistetyt käänteiskuvaukset	27
2.1	Lineaarista operaattoreista	27
2.2	Singulaariarvohajotelma	39
2.3	Pseudoinverssi	43
3	Tihonov-regularisointi	54
3.1	Pseudoinverssin jatkuvuus	54



Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

3.2	SVD-muoto	56
3.3	Konvergenssi ja kohinainen data	58
3.4	Variaatiomuoto	61
3.5	Yleistetty Tihonov-regularisointi	63
3.6	Sivurajoitteen merkitys	66
3.7	Epälineaariset käänteisongelmat	70
3.8	Regularisointiparametrin valinta	73
3.9	Tihonov-regularisointi ja normaalijakaumaoletukset	76

4	Iteratiivisia menetelmiä	87
4.1	Kiintopisteiteraatiot	87
4.2	Landweber-iteraatiot	90
4.3	Konjugaattigradientti-iteraatiot	93
4.4	Kaczmarz-iteraatiot	97
4.5	Esimerkki iteratiivisista menetelmistä	104

Etusivu



Sivu 3 / 108

Takaisin

Full Screen

Sulje

Poistu

**Sisältö:**

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 4 / 108

Takaisin

Full Screen

Sulje

Poistu

1.1. Suorat ongelmat ja käänteisongelmat

Käänteisongelmien yksikäsitteinen määrittely on valitettavasti mahdotonta. Suoralla ongelmalla tässä yhteydessä tarkoitetaan seuraavantyyppistä ongelmaa. Olkoon f “syy” ja g “seuraus” ja K näitä yhdistävä matemaattinen malli siten, että $K(f) = g$. Jos f ja K tunnetaan ja K on *hyvin asetettu* (*Hadamardin mielessä*), vastaa jokaista syytä f yksikäsitteinen seuraus g ja pienet muutokset syyssä f aiheuttavat pienen muutoksen seurauksessa g . Tähän suoraan ongelmaan liittyen on olemassa kahden tyyppisiä *käänteisongelmia*:

- Olkoon g havaittu seuraus ja oletetaan, että malli K tunnetaan. Tehtävänä on määrätä syy f . Tällainen tapaus on esimerkiksi parabolisen osittaisdifferentiaaliyhtälön alkuehdon määrittäminen, kun ratkaisu on havaittu jollakin myöhemmällä hetkellä. Tällainen ongelma on myös sädehoidon annossuunnittelu, jossa haluttu säteilyn absorbtiojakauma sekä arvio väliaineen absorbtio- ja sirontaominaisuuksille tunnetaan. Tehtävänä on määrätä, kuinka väliainetta pitää “säteilyttää” eri suunnista, jotta haluttu absorbtiojakauma saavutetaan mahdollisimman tarkasti.



- Olkoon g havaittu seuraus ja f syy. Tehtävänä on nyt määrätä näitä yhdistävä matemaattinen malli K . Käytännössä tämä tarkoittaa usein sitä, että mallin “parametrinen” muoto tunnetaan, mutta jotkin näistä parametreista ovat tuntemattomia. Tällaisia ongelmia ovat tomografiaongelmat, joista esimerkiksi röntgentomografiassa väliaineeseen “syötetään” röntgensäteitä ja havaitaan kappaleen toisella puolella transmittoituneet intensiteetit, jotka riippuvat väliaineen absorptio- ja sirontaominaisuuksista. Siten tämä ongelma on tavallaan sädehoidon annossuunnittelulle “komplementaarinen” ongelma.

1.1.1. Inverse crimes

Käsitteellä *inverse crime* viitataan siihen, että simuloinneissa data g (suora ongelma) lasketaan samalla diskreetoinnilla (operaattorin K realisoinnilla), jota käytetään vastaavan käänteisongelman ratkaisemisessa. Tämä valinta tuottaa lähes poikkeuksetta todellisuuteen verrattuna liian hyviä arvioita kyseisen käänteisongelman ratkeavuudesta. Tämä koskee erityisesti kohinattomia tilanteita. Joissakin tapauksissa tämä virhe voidaan välttää tarkastelemalla kohinaista dataa. Ongelmaksi jää tällöin, kuinka paljon kohinaa (ja millä rakenteella) dataan tulee vähintään lisätä. Tässä kirjassa syyllystään muutamassa esimerkissä kyseiseen virheeseen. Syynä on tällöin ollut ohjelmallisten toteutusten seurattavuus.

1.2. Esimerkkejä suorista ongelmista

1.2.1. CT-havaintomalli

Tarkastellaan äärettömän kapeaa röntgensädettä, joka kulkee siroamatta kohteen Ω läpi. Kohteessa säteen intensiteetti I noudattaa vaimennuslakia

$$\frac{dI}{ds} = -f(x)I,$$

jossa $f(x)$ on massavaimennuskerroin pisteessä $x \in \mathbb{R}^2$.

Ratkaisemalla yhtälö I :n suhteen saadaan

$$I = I_0 \exp\left(-\int_{\gamma_\ell} f(x) ds\right),$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



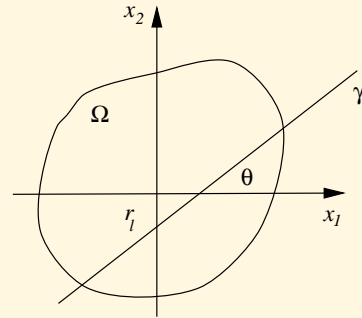
Sivu 5 / 108

Takaisin

Full Screen

Sulje

Poistu



Kuva 1.1: Viivaintegraalit pitkin käyriä $\gamma_\ell : x_2 - r_\ell = \theta_\ell x_1$.

jossa I_0 on intensiteetti ennen kohdetta (incident beam), I on intensiteetti kohteen jälkeen ja $\int_{\gamma_\ell} f(x) ds$ on viivaintegraali pitkin sädettä γ_ℓ . On täysin riittävää tarkastella suhteen I_0/I logaritmia, jolloin ongelma voidaan kirjoittaa muotoon

$$\ln \frac{I_0}{I} = g_\ell = \int_{\gamma_\ell} f(x) ds . \quad (1.1)$$

Yhteen säteeseen liittyvä suoran ongelman ratkaisu on siten tämän viivaintegraalin laskeminen, kun tiheysjakauma $f(x)$ tunnetaan, kuva 1.1.

Suoran ongelman numeerinen ratkaiseminen voidaan luonnollisesti tehdä usealla tavalla. Oleellista käänteisongelmien yhteydessä on kuitenkin, että suora ongelma esitetään operaattoriyhtälömuodossa

$$g_\ell = K_\ell(f) .$$

Jaetaan Ω pistevieraisiin osajoukkoihin (tyypillisesti taso jaetaan neliöihin tai kolmioihin) Ω_m ,

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 6 / 108

Takaisin

Full Screen

Sulje

Poistu



jotka peittävät joukon Ω . Siis

$$\bigcup_{m=1}^M \Omega_m = \Omega ,$$
$$\int_{\Omega} \chi_m(x) \chi_k(x) = 0 , \quad m \neq k \quad \text{ja}$$
$$\chi_m(x) = \begin{cases} 1, & x \in \Omega_m \\ 0, & \text{muulloin,} \end{cases}$$

missä $\chi_m(x)$ on joukon Ω_m indikaattorifunktio. Approksimoidaan nyt funktiota $f(x)$ funktiolla, joka on paloittain vakio siten, että

$$f(x) \approx \tilde{f}(x) = \sum_{m=1}^M f_m \chi_m(x) ,$$

missä f_m on massavaimennuskerroin elementissä Ω_m . Tällöin integraali (1.1) voidaan kirjoittaa muotoon

$$\int_{\gamma_\ell} f(x) \approx \int_{\gamma_\ell} \tilde{f}(x) = \sum_{m=1}^M K_\ell(m) f_m ,$$

missä $K_\ell(m)$ on suoran γ_ℓ elementtiä Ω_m leikkaavan janan pituus, kuva 1.2.

Kun havaitaan vaimenemat $-g_\ell$, $\ell = 1, \dots, N$ (N on säteiden lukumäärä), useilta suorilta, saadaan havaintoyhtälöksi

$$g = Kf ,$$
$$g = (g_1, \dots, g_N)^T ,$$
$$f = (f_1, \dots, f_M)^T ,$$
$$K_\ell = (K_\ell(1), \dots, K_\ell(M)) \quad \text{ja}$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 7 / 108

Takaisin

Full Screen

Sulje

Poistu



Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Ω_1	Ω_2	Ω_3
Ω_4	Ω_5	Ω_6
Ω_7	Ω_8	Ω_9

γ_ℓ

$$K_\ell = [0 \ 0 \ 0 \ 0 \ 0 \ 1/3 \ 0 \ 1/3 \ 1]$$
$$g_\ell \approx \sum_{m=1}^M K_\ell(m) f_m$$

Kuva 1.2: Suoran ongelman approksimointi röntgentomografiassa.

$$K = \begin{pmatrix} K_1 \\ \vdots \\ K_N \end{pmatrix} = \begin{pmatrix} K_1(1) & \cdots & K_1(M) \\ \vdots & \ddots & \vdots \\ K_N(1) & \cdots & K_N(M) \end{pmatrix}.$$

Matriisi K voi olla alideterministinen, eli $M > N$, sekä lisäksi useissa sovelluksissa hyvin suuri, mutta usein myös hyvin harva matriisi, kuva 1.3.

Yleensä CT-kuvantamisen yhteydessä datan laatu on hyvä, eli havainnot g voidaan katsoa kohinattomiksi intensiteettimittausten osalta. Sen sijaan karkeasta diskretoinnista saattaa aiheutua merkittäviä mallinnusvirheitä, joita yleensä käsitellään additiivisina, mikä sinänsä on virheellinen menettely.

Etusivu



Sivu 8 / 108

Takaisin

Full Screen

Sulje

Poistu



Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



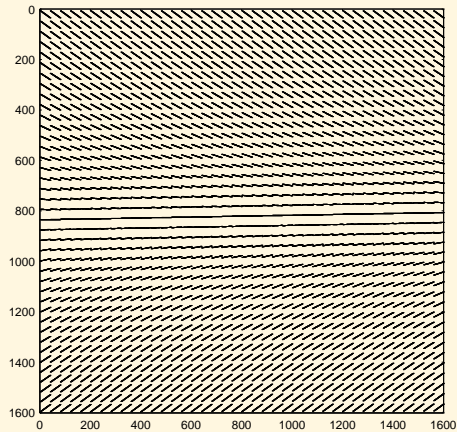
Sivu 9 / 108

Takaisin

Full Screen

Sulje

Poistu



Kuva 1.3: Esimerkki kuvausmatriisiin K rakenteesta: nolasta eroavat alkio, 1600 sädetä: 40 siirrosta, 40 kulmaa. Alue Ω on jaettu 40×40 neliöön Ω_k . Matriisin K tiheys on tässä tapauksessa luokkaa 0.03.



1.2.2. Alipäästösuodatus/dekonvoluutio

Klassinen esimerkki kuvantamiseen liittyvästä käänteisongelmasta on konvoluutiotyypinen alipäästösuodatus. Tarkastellaan seuraavassa Fredholmien ensimmäisen lajin integraaliyhtälöä

$$g(x) = \int_{\Omega} K(x, y) f(y) dy, \quad x \in \mathbb{R}^2, \quad \Omega \subset \mathbb{R}^2, \quad (1.2)$$

missä K on integraaliyhtälön ydin, $x = (x_1, x_2)$ ja $y = (y_1, y_2)$. Tavoitteena on rekonstruoida kuva $f(x)$ havaitun vääristyneen kohinaisen kuvan $g(x)$ perusteella. Olkoon $\Omega = [0, 1] \times [0, 1]$ ja

$$K(x, y) = \exp(-\kappa \|x - y\|), \quad (1.3)$$

missä $\kappa > 0$. Ongelman diskretisoimiseksi Ω jaetaan elementteihin (neliöihin) Ω_j , $j = 1, \dots, N$, ja kuvaa f approksimoidaan funktiolla, joka on vakio kussakin elementissä. Tällöin saadaan äärellisdimensioinen havaintomalli $g = Kf + e$, missä $g_j = g(x^{(j)})$ on havainto kuva-alkion (pikselin) keskikohdassa $x^{(j)}$, e on additiivinen havaintovirhe ja

$$K_{i,j} = |\Omega_j| \exp(-\kappa \|x^{(i)} - y^{(j)}\|),$$

missä $|\Omega_j|$ on elementin Ω_j mitta (pinta-ala). Kuvassa 1.4 on esitetty esimerkki kohteesta, havainnosta ja eräästä ratkaisusta.

1.2.3. Sirontaongelma

Yksi klassisista sirontaan liittyvistä käänteisongelmista on määrätä kappaleen muoto perustuen tästä kappaleesta sironneen sähkömagneettisen tai akustisen aallon mittauksiin kaukana kohteesta.

Ajan suhteen harmoninen aaltoliike toteuttaa sirottavan kappaleen Ω ulkopuolella Helmholtzin yhtälön

$$\Delta u(x) + k^2 u(x) = 0, \quad x \in \mathbb{R}^2 \setminus \Omega. \quad (1.4)$$

Kokonaiskenttä voidaan jakaa tulevaan (tasoaalto) ja sironneeseen kenttään

$$u(x) = u_i(x) + u_s(x) = e^{ikx_1} + u_s(x), \quad x = (x_1, x_2) \in \mathbb{R}^2, \quad (1.5)$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 10 / 108

Takaisin

Full Screen

Sulje

Poistu



Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 11 / 108

Takaisin

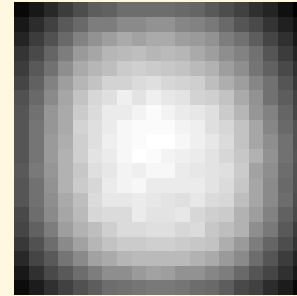
Full Screen

Sulje

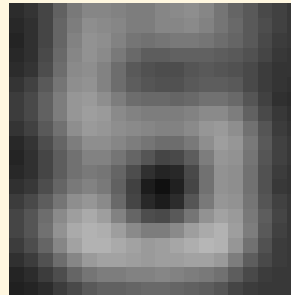
Poistu



a)



b)



c)

Kuva 1.4: a) Todellinen objekti $f(x)$, b) havainto $g(x)$ ja c) tyypillinen regularisoitu ratkaisu sileysoletuksilla. Kohinan taso on 2% datan "intensiteetistä" ja $\kappa = 4$.



missä k on aaltoluku, u_i (tunnettu suuntaan x_1 etenevä) tuleva aalto ja u_s sironnut aalto. Kaukana kohteesta sironneelle kentälle approksimoidaan Sommerfeldin säteilyehdon perusteella

$$\frac{\partial u_s(x)}{\partial r} - iku_s(x) \approx 0. \quad (1.6)$$

Sirottavan kappaleen pinnalla reunaehto voi olla jokin seuraavista (n on pinnan normaali):

- sound-soft kappale: $u(x) = 0$, $x \in \partial\Omega$ (Dirichlet),
- sound-hard kappale: $\frac{\partial u(x)}{\partial n} = 0$, $x \in \partial\Omega$ (Neumann) ja
- muulloin: $\frac{\partial u(x)}{\partial n} + i\lambda u(x) = 0$, $x \in \partial\Omega$ (Robin).

Sirontaongelmaan liittyvä suora ongelma on määrätä sironnut kenttä kaukana kohteesta, kun kappaleen muoto ja reunaehto tunnetaan. Kun suoraa ongelmaa approksimoidaan äärellisten elementtien menetelmällä, päädytään matriisiyhtälöön

$$(G(f) + R(f))u^h = F(f),$$

missä $G(f)$ ja $R(f)$ ovat Galerkinin menetelmään liittyvät jäykkyys- ja massamatriisit, joiden alkiot riippuvat suureesta f , vektori F riippuu muun muassa reunaehdoista ja u^h on kentän approksimaatio hilapisteissä. Suure f on (parametrinen) esitysmuoto kappaleen reunalle ja mahdollisesti myös joillekin muille parametreille. Kuvaus $f \rightarrow G(f)$ on lineaarinen.

Muodollisesti saadaan

$$u^h = (G(f) + R(f))^{-1} F(f) = \tilde{K}(f), \quad (1.7)$$

jonka kuvaus on epälineaarinen. Olkoon nyt

$$g = u^h|_{\partial\Omega'} = Qu^h,$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 12 / 108

Takaisin

Full Screen

Sulje

Poistu



missä $\partial\Omega'$ on laskenta-alueen reuna ja Q on operaattori (matriisi), joka “palauttaa” FEM-approksimaatiosta u^h kentän niissä pisteissä, joissa mittaukset on suoritettu, tässä tapauksessa siis laskenta-alueen reunalla. Siten suora ongelma saadaan muotoon

$$\begin{aligned} g &= Q(G(f) + R(f))^{-1}F(f) \\ &\doteq K(f) . \end{aligned}$$

Lisäksi ratkaistaessa suuretta f tulee esimerkiksi Gauss-Newton -menetelmää käytettäessä ratkaista yllä olevan kuvauksen Jacobin matriisi J , jolloin ongelmaa approksimoidaan sekventiaalisesti muodolla

$$g \approx K(f_0) + J(f_0)(f - f_0) .$$

1.3. Ratkaisuhilan valinta ja derivointi

Eräs keskeisistä reguloitintavoista on laskennassa käytettävän approksimaation diskretoinnin säätäminen. Olkoon $f(x) = \sin n\pi x$, $x \in [0, 1] \subset \mathbb{R}$ ja $n \in \mathbb{N}$. Tämän integraalin ylärajan funktio on $g(x) = \int_0^x f(y) dy = (n\pi)^{-1}(1 - \cos n\pi x)$, eli

$$\begin{aligned} g(x) &= \int_0^x f(y) dy = \int_0^1 K(x, y) f(y) dy \Leftrightarrow \\ g &= Kf , \end{aligned}$$

missä kyseisen *ensimmäisen lajin Fredholm'n integraaliyhtälön* ydin $K(x, y)$ on

$$K(x, y) = \begin{cases} 1 & , y \leq x \\ 0 & , y > x \end{cases} .$$

Luonnollisesti mikäli $g(x)$ tunnetaan, ja tavoitteena on määrätä se funktio $f(x)$, jonka integraalifunktio $g(x)$ on, derivoidaan $g(x)$, eli

$$f(x) = g'(x) = K^{-1}g .$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 13 / 108

Takaisin

Full Screen

Sulje

Poistu



Tarkastellaan kuitenkin tilannetta, jossa funktion $g(x)$ sijasta havaitaan tämä funktio pisteissä $x_1 = 0, \dots, x_N = 1$. Lisäksi havaintoihin on summautunut (normaalijakautunutta) kohinaa, joten havainnot ovat muotoa

$$g_k^\delta \doteq g(x_k) + e_k = g_k + e_k,$$

missä $E\{e_k\} = 0$ (odotusarvo), $\text{var}(e_k) = \delta^2$ (varianssi) ja $E\{e_k e_m\} = 0, k \neq m$, eli virheet e_k ovat toisistaan riippumattomia. Korvataan operaattori K^{-1} differenssioperaattorilla

$$Dg_k = \frac{g_k - g_{k-1}}{x_k - x_{k-1}}$$

ja tarkastellaan näin saatavaa approksimaatiota f_ℓ^δ samassa tasavälisessä hilassa $y_\ell = x_k$ (hilavakio $h = x_k - x_{k-1}$). Nyt saadaan

$$\begin{aligned} f_k^\delta &\doteq Dg_k^\delta = D(g_k + e_k) = Dg_k + De_k \\ &= \frac{g_k - g_{k-1}}{x_k - x_{k-1}} + \frac{e_k - e_{k-1}}{x_k - x_{k-1}}. \end{aligned}$$

Näistä ensimmäinen termi $\rightarrow g'(x_k) = f(x_k)$, kun $x_k - x_{k-1} \rightarrow 0$. Toinen termi edustaa siten funktion $f(x)$ estimointivirhettä. Tämä termi on satunnainen, joten määrätään sen varianssi

$$\begin{aligned} \text{var} \left(\frac{e_k - e_{k-1}}{x_k - x_{k-1}} \right) &= E \left\{ \left(\frac{e_k - e_{k-1}}{x_k - x_{k-1}} \right)^2 \right\} - E \left\{ \frac{e_k - e_{k-1}}{x_k - x_{k-1}} \right\}^2 \\ &= (x_k - x_{k-1})^{-2} E \{ e_k^2 + e_{k-1}^2 - 2e_k e_{k-1} \} \\ &= (x_k - x_{k-1})^{-2} (E \{ e_k^2 \} + E \{ e_{k-1}^2 \} - 2E \{ e_k e_{k-1} \}) \\ &= \frac{2\delta^2}{(x_k - x_{k-1})^2} \\ &\rightarrow \infty, \quad \text{kun } x_k - x_{k-1} \rightarrow 0. \end{aligned}$$

Siten estimaatin virheen varianssi on sitä suurempi, mitä *tarkemmassa* hilassa laskenta suoritetaan. Kuvassa 1.5 on esitetty havainnot g_h^δ , todellinen integraalifunktio $g(x)$, $f(x)$ sekä estimaatit

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 14 / 108

Takaisin

Full Screen

Sulje

Poistu



f_h^δ joillakin hilavakion $h = x_k - x_{k-1}$ arvoilla. Oheisen **Matlab esimerkin** avulla voit myös itse säätää lisättävän kohinan ja hilavakion arvoa.

Edellä kuvattu ongelma voidaan esittää matriisi muodossa $g = Kf$, kun $g = (g_1, \dots, g_N)$, $f = (f_1, \dots, f_N)$, $f_k = f(x_k)$ ja matriisi

$$K = h \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 1 & 1 & \cdots & 1 \end{pmatrix}$$

jolloin

$$K^{-1} = h^{-1} \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -1 & 1 \end{pmatrix}.$$

On selvää, että esimerkiksi normin $\|f - f_k^\delta\|_2$ mielessä hilavakion h “optimaalinen” arvo riippuu kohinan “suuruudesta” δ . Tässä esimerkissä ongelman regularisointi suoritettiin ratkaisuhilan diskretointia säätämällä. Tämä on perinteinen lähestymistapa, jota ei nykyään usein suositella.

1.4. Katkaistut hajoitelmat ja “terminen arkeologia”

Toinen keskeinen perinteinen regulointimenetelmä on ratkaisujen sarjamuotojen katkaiseminen. Tarkastellaan tilannetta, jossa yksidimensioisen metallitangon $x \in [0, \pi]$ lämpötilajakauma $g(x, t)$ hetkellä $t = 0$ on $g(x, 0) = f(x)$. Tangon molemmat päät pidetään lämpötilassa $g(0, t) = g(\pi, t) = 0$ kaikille t . Havaitaan tangon lämpötilajakauma $g(x, T)$. Tehtävänä on nyt tämän havainnon perusteella arvioida tangon lämpötilan alkujakaumaa $f(x)$. Ongelma on hankala, koska

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 15 / 108

Takaisin

Full Screen

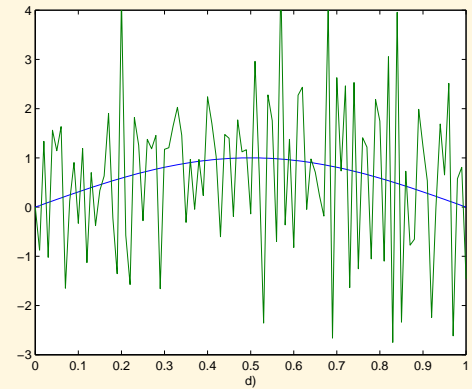
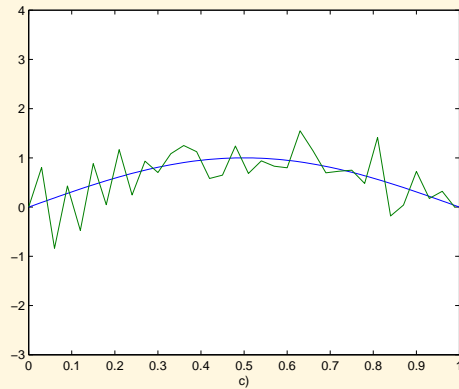
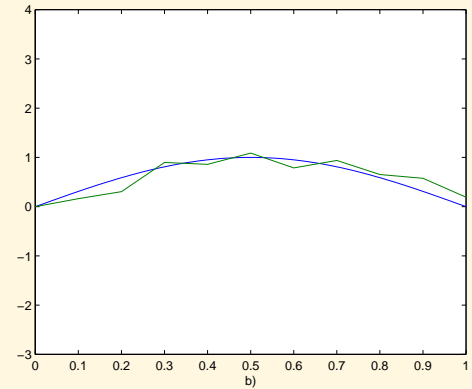
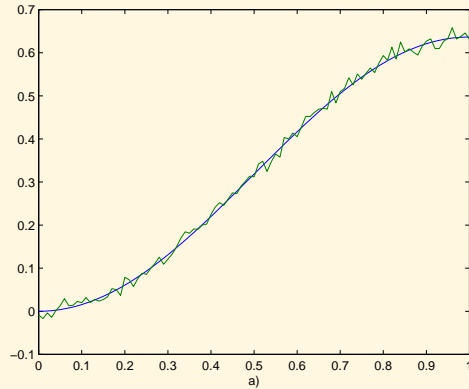
Sulje

Poistu



Sisältö:

1. Johdanto
2. Yleistetyt käännteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä



Kuva 1.5: a) Integraalifunktio $g(x)$ ja kohinaiset havainnot g_h^δ , b)-d) $f(x)$ ja estimaatit f_h^δ , kun $h = 0.1, 0.03, 0.01$, vastaavasti. [\[KOODI\]](#)

Etusivu



Sivu 16 / 108

Takaisin

Full Screen

Sulje

Poistu



hyvinkin erillaiset alkujakaumat tasaantuvat samanmuotoiseksi jakaumaksi hyvinkin nopeasti, jolloin kohinaisista havainnoista on hankala päätellä alkujakauman muotoa. Katso kuva 1.6 ja [animaatio lämpötilan muuttumisesta](#).

Tarkasteltavaa tilannetta kuvaa lämpöyhtälö (parabolinen osittaisdifferentiaaliyhtälö)

$$\frac{\partial g}{\partial t} = \nabla \cdot (\gamma \nabla g) .$$

Tarkastellaan yksiulotteista, normeerattua (diffuusiovakio $\gamma \equiv 1$) tilannetta, jolloin lämpöyhtälö voidaan kirjoittaa muotoon

$$\frac{\partial g}{\partial t} = \frac{\partial^2 g}{\partial x^2} . \quad (1.8)$$

Kirjoitetaan ODY:n ratkaisu Fourier-sarjana (Laplace-operaattorin [ominaisfunktioiden](#) avulla)

$$g(x, t) = \sum_{n=1}^{\infty} a_n(t) \sin nx \quad (1.9)$$

Sijoittamalla edellinen yhtälöön (1.8) saadaan

$$\sum_{n=1}^{\infty} a'_n(t) \sin nx = - \sum_{n=1}^{\infty} a_n(t) n^2 \sin nx .$$

Lineaarisen riippumattomuuden perusteella ajasta riippuville kertoimille saadaan

$$a'_n(t) = -n^2 a_n(t) , \quad n \geq 1. \quad (1.10)$$

Edelleen hetkellä $t = 0$

$$f(x) = \sum_{n=1}^{\infty} a_n(0) \sin nx ,$$

jolloin [saadaan](#)

$$a_n(0) = \frac{2}{\pi} \int_0^{\pi} f(s) \sin ns \, ds .$$

Sisältö:

1. [Johdanto](#)
2. [Yleistetyt käänneiskuvaukset](#)
3. [Tihonov-regularisointi](#)
4. [Iteratiiviset menetelmät](#)
5. [Tehtäviä](#)

Etusivu



Sivu 17 / 108

Takaisin

Full Screen

Sulje

Poistu



Näin ollen yhtälöiden (1.10) ratkaisut ovat

$$\begin{aligned} a_n(t) &= a_n(0)e^{-n^2t} \\ &= \frac{2}{\pi} \int_0^\pi f(s) \sin ns \, ds e^{-n^2t} . \end{aligned}$$

Nähdään, että alkutilanteessa $f(x)$ mahdollisesti olleet, suuria spatiaalisia taajuuksia (n) edustaneet komponentit $\sin nx$ häviävät nopeasti. Kuvassa 1.6 on esitetty lämpötilajakauman evoluutio erään alkujakauman tapauksessa.

Tarkastellaan seuraavaa ongelmaa. Olkoon $g(x, T)$ tangon lämpötila ajanhetkellä $t = T$ (havainto). Tavoitteena on määrätä aluksi kuvaus $f \mapsto g(\cdot, T)$ ja tämän jälkeen tutkia käänteisen ongelman $g(\cdot, T) \mapsto f$ ratkeavuutta. Jakauma hetkellä $t = T$ saadaan nyt muotoon

$$\begin{aligned} g(x, T) &= \sum_{n=1}^{\infty} \frac{2}{\pi} \int_0^\pi f(s) \sin ns \, ds e^{-n^2T} \sin nx \\ &= \int_0^\pi K(x, s) f(s) \, ds \quad \text{ja} \end{aligned} \tag{1.11}$$

$$K(x, s) = \frac{2}{\pi} \sum_{n=1}^{\infty} e^{-n^2T} \sin ns \sin nx , \tag{1.12}$$

missä yhtälön (1.11) perusteella ongelma on jälleen *ensimmäisen lajin Fredholmian integraaliyhtälö*.

Tarkastellaan nyt hetkeä $T = 1$ ja merkitään

$$\begin{aligned} v_k &= v_k(x) = \sin kx , \\ u_k &= u_k(x) = \sin kx , \\ \sigma_k &= e^{-k^2} \quad \text{ja} \\ g &= g(x) = g(x, T) , \end{aligned}$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 18 / 108

Takaisin

Full Screen

Sulje

Poistu



Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



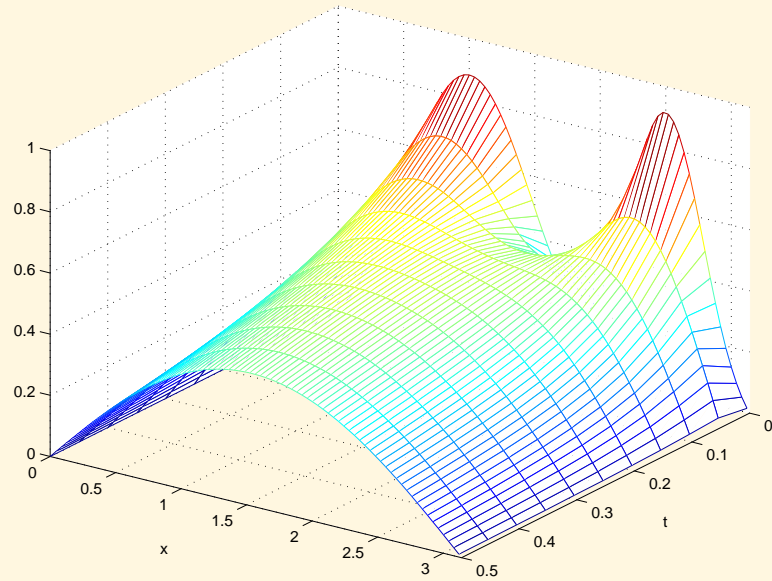
Sivu 19 / 108

Takaisin

Full Screen

Sulje

Poistu



Kuva 1.6: Lämpötilajakauman evoluutio erään alkujakauman tapauksessa. [\[KOODI\]](#)



missä funktiot v_k ja u_k eivät ole kaikissa tapauksissa samoja. Määritellään sisätulo

$$\begin{aligned}\langle f, v_k \rangle &= \frac{2}{\pi} \int_0^\pi f(s) v_k(s) ds \\ &= \frac{2}{\pi} \int_0^\pi f(s) \sin ks ds ,\end{aligned}$$

jolloin $\langle v_k, v_n \rangle = \langle u_k, u_n \rangle = \frac{2}{\pi} \int_0^\pi \sin kx \sin nx dx = \delta_{kn}$, eli joukot $\{v_k\}$ ja $\{u_k\}$ ovat ortonormaaleja. Suoran ongelman ratkaisu saadaan muotoon

$$g = \sum_{k=1}^{\infty} \sigma_k \langle f, v_k \rangle u_k . \quad (1.13)$$

Kirjoitetaan f ja g Fourier-sarjamuotoon

$$\begin{aligned}f &= \sum_{k=1}^{\infty} \langle f, v_k \rangle v_k \quad \text{ja} \\ g &= \sum_{k=1}^{\infty} \langle g, u_k \rangle u_k ,\end{aligned} \quad (1.14)$$

mistä (1.13) ja (1.14) avulla saadaan (sinin ortogonaalisuus)

$$\begin{aligned}\langle g, u_k \rangle &= \sigma_k \langle f, v_k \rangle \quad \text{ja} \\ f &= \sum_{k=1}^{\infty} \frac{\langle g, u_k \rangle}{\sigma_k} v_k .\end{aligned} \quad (1.15)$$

Oletetaan, että havaittu $g(x)$ on tarkka (kohinaton) ja että tehtävänä on saada alkuarvolle $f(x)$ numeerinen ratkaisu. Tätä varten diskretoidaan ominaisfunktiot tasaväliseen hilaan

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 20 / 108

Takaisin

Full Screen

Sulje

Poistu



$x_1 = 0, \dots, x_N = \pi$ ja katkaistaan sarja (1.15) termiin N . Jatkossa sekä funktioita että näistä diskretoinnilla saatuja näytevektoreita merkitään samoilla symboleilla. Kirjoitetaan

$$\begin{aligned} V = U &= [\sin x, \dots, \sin Nx] \text{ ja} \\ S &= \text{diag}(e^{-1}, \dots, e^{-N^2}), \end{aligned}$$

missä matriisit U ja V ortonormeerataan diskretoinnin jälkeen (nämä matriisit ovat jo lähes ortogonaalisia). Nyt (laskuharjoitukset)

$$\begin{aligned} g = USV^T f &\Leftrightarrow U^T g = SV^T f \\ &\Leftrightarrow f = VS^{-1}U^T g, \end{aligned} \quad (1.16)$$

kuten edellä. Olkoon laskentaan käytettävän tietokoneen “kone-epsilon” $\epsilon = 10^{-16}$. Kuinka suuri voi N olla, jotta suureen S^{-1} kanssa ei tule ongelmia? Konservatiivinen arvio tälle saadaan määräämällä $\exp(-N^2) \approx \epsilon$, jolloin $N \approx \sqrt{-\log \epsilon} \approx 6$. Siten kyseisestä sarjaratkaisusta voitaisiin tietokoneella laskea vain kuusi ensimmäistä termiä, jos sisätulot $U^T g$ eivät ole hyvin pieniä.

Osoittautuu siis, että matriisia $VS^{-1}U^T$ ei numeerisista syistä voida muodostaa kuin hyvin pienille N . Kuvassa 1.7 on esitetty hetkellä $T = 0.05$ havaittu jakauma ja alkujakauman kolme rekonstruktioita, jotka perustuvat muodon (1.15) katkaisemiseen. Tässä tapauksessa huomattavia virheitä indusoiuu, kun termejä on yli 14.

Edellisessä esimerkissä ongelman stabiilisuutta säädettiin diskreetointihilan avulla. Toinen klassinen regularisointitapa on pitää ratkaisuhila suhteellisen tiheänä, mutta katkaista esitysmuoto (1.15) seuraavasti

$$\hat{f}_r = \sum_{k=1}^r \frac{\langle g, u_k \rangle}{\sigma_k} v_k. \quad (1.17)$$

Tämä katkaaisu aiheuttaa ratkaisuun ainakin vastaavan approksimaatiovirheen

$$\tilde{f}_r = \sum_{k=r+1}^{\infty} \frac{\langle g, u_k \rangle}{\sigma_k} v_k, \quad (1.18)$$

$$(1.19)$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 21 / 108

Takaisin

Full Screen

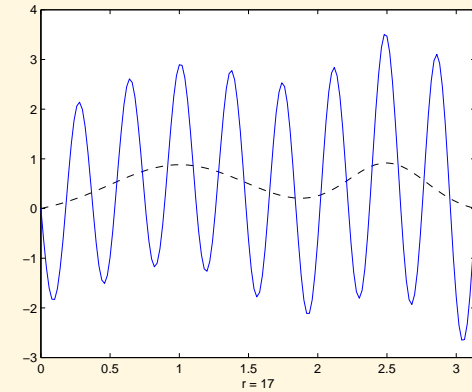
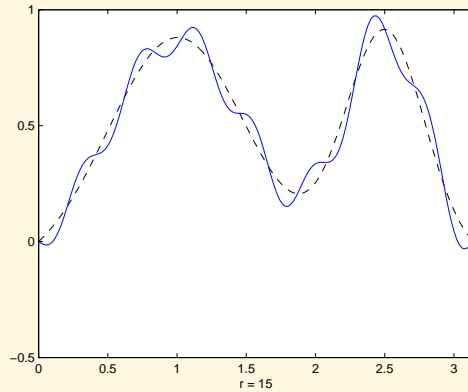
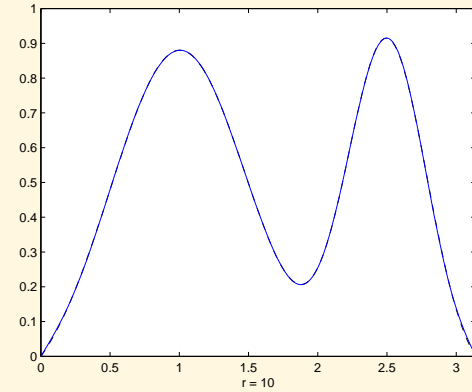
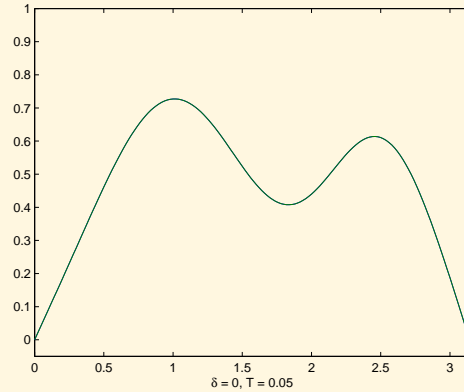
Sulje

Poistu



Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä



Kuva 1.7: a) Havainto $g(x, T = 0.05)$, b)-c) alkujakauman $f(x)$ rekonstruktio, kun summasta (1.15) on otettu vastaavasti mukaan $r = 10, 15$ ja 17 termiä. Katkoviivalla esitetty oikea ratkaisu.

[\[KOODI\]](#)

Etusivu



Sivu 22 / 108

Takaisin

Full Screen

Sulje

Poistu



jolloin

$$\|\tilde{f}_r\|^2 = \sum_{k=r+1}^{\infty} |\langle f, v_k \rangle|^2. \quad (1.20)$$

Tarkastellaan seuraavaksi tilannetta, jossa hetkellä $T = 1$ havaitaan “kohinainen” jakauma

$$g^\delta = g + e,$$

missä e on satunnaisvektori, jolle $E\{e_k\} = 0$, $\text{var}(e_k) = \delta^2$ kaikille k ja alkiot ovat toisistaan riippumattomia, eli $E\{e_k e_\ell\} = 0$, kun $k \neq \ell$. Nyt ratkaisulle saadaan sisätulon bilineaarisuuden nojalla

$$\begin{aligned} f^\delta &= \sum_{k=1}^r \frac{\langle g, u_k \rangle}{\sigma_k} v_k + \sum_{k=1}^r \frac{\langle e, u_k \rangle}{\sigma_k} v_k \\ &= f + \tilde{f}. \end{aligned}$$

Koska $\{v_k\}$ ja $\{u_k\}$ ovat ortonormaaleja, saadaan edelleen

$$\begin{aligned} \|\tilde{f}\|^2 = \tilde{f}^T \tilde{f} &= \left\langle \sum_{k=1}^r \frac{\langle e, u_k \rangle}{\sigma_k} v_k, \sum_{n=1}^r \frac{\langle e, u_n \rangle}{\sigma_n} v_n \right\rangle \\ &= \sum_{k=1}^r \sum_{n=1}^r \frac{\langle e, u_k \rangle}{\sigma_k} \frac{\langle e, u_n \rangle}{\sigma_n} \langle v_k, v_n \rangle \\ &= \sum_{k=1}^r \frac{\langle e, u_k \rangle^2}{\sigma_k^2}. \end{aligned}$$

Sisätulot $\langle x, y \rangle$, $x, y \in \mathbb{R}^N$, voidaan kirjoittaa muodossa $x^T y$, jolloin $\langle e, u_k \rangle^2 = u_k^T e e^T u_k$ ja $E\{e e^T\} = \delta^2 I$. Näin ollen koska $\{u_k\}$ ovat ortonormaaleja sekä $\sigma_k^{-2} = (e^{-k^2})^{-2} = e^{2k^2}$, saadaan,

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 23 / 108

Takaisin

Full Screen

Sulje

Poistu



että

$$\begin{aligned} E \left\{ \|\tilde{f}\|^2 \right\} &= E \left\{ \sum_{k=1}^r \frac{\langle e, u_k \rangle^2}{\sigma_k^2} \right\} \\ &= \sum_{k=1}^r \frac{u_k^T E \{ e e^T \} u_k}{\sigma_k^2} \\ &= \sum_{k=1}^r \frac{u_k^T \delta^2 I u_k}{\sigma_k^2} \\ &= \delta^2 \sum_{k=1}^r \frac{u_k^T u_k}{\sigma_k^2} \\ &= \delta^2 \sum_{k=1}^r \sigma_k^{-2} \\ &= \delta^2 (e^{2 \cdot 1} + e^{2 \cdot 2^2} + \dots + e^{2 \cdot k^2} + \dots + e^{2 \cdot r^2}), \end{aligned}$$

josta nähdään selvästi, kuinka häiriö voimistuu nopeasti lisättäessä termejä summaan. Toisaalta approksimaatiovirhe pienenee lisättäessä näitä termejä. Fourier-sarjojen teoriasta tiedetään (tietyillä lisäoletuksilla), että jos

$$\begin{aligned} f \text{ on jatkuva, niin} & \quad |\langle f, v_k \rangle| \leq c_0 k^{-1}, \\ f \text{ on jatkuvasti derivoituva, niin} & \quad |\langle f, v_k \rangle| \leq c_1 k^{-2} \quad \text{jne.} \end{aligned}$$

Mitä sileämpi jakauman $f(x)$ oletetaan olevan, sitä vähemmän hajotelman termejä voidaan käyttää kiinteällä approksimaatiovirheellä. Siten nämä oletukset vaikuttavat estimoinnissa tehtäviin valintoihin.

Kuvissa 1.8 ja 1.9 on esitetty kaksi tapausta ($T = 0.25$ ja $T = 0.5$, $\delta = 0.02$) kohinaisista havainnoista sekä alkujakauman kolme rekonstruktiota, jotka perustuvat muodon (1.15) katkaisemiseen.

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 24 / 108

Takaisin

Full Screen

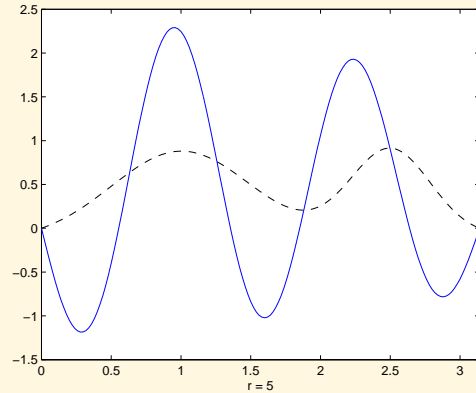
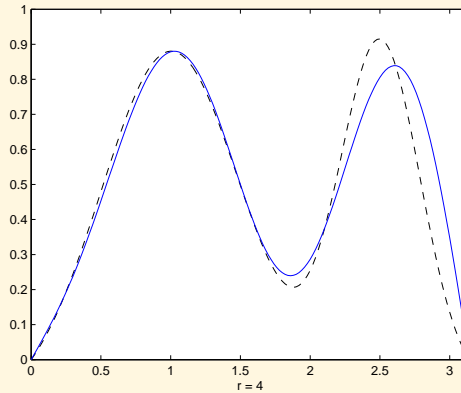
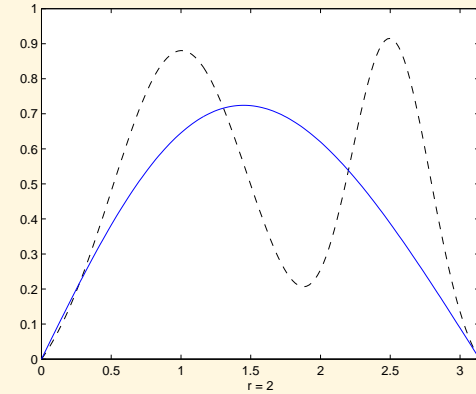
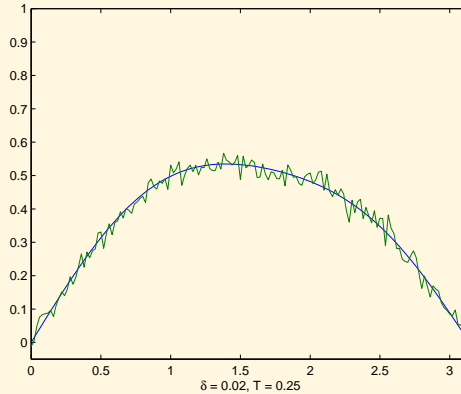
Sulje

Poistu



Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä



Kuva 1.8: a) Havainto $g(x, T = 0.25)$, b)-c) alkujakauman $f(x)$ rekonstruktio, kun summasta (1.15) on otettu vastaavasti mukaan $r = 2, 4$ ja 5 termiä. Katkiviivalla esitetty oikea ratkaisu. [\[KOODI\]](#)

Etusivu



Sivu 25 / 108

Takaisin

Full Screen

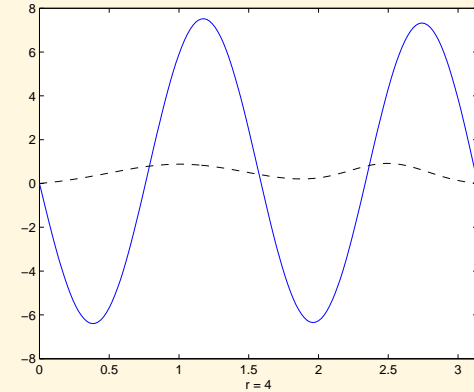
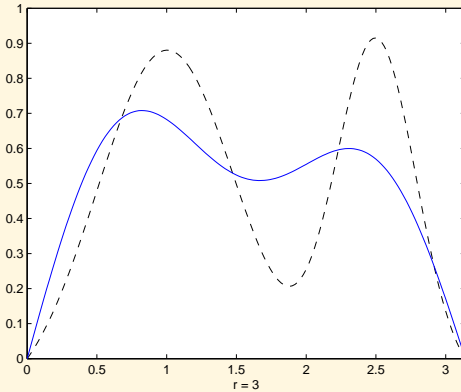
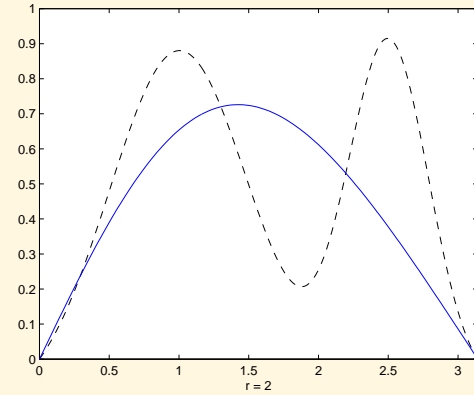
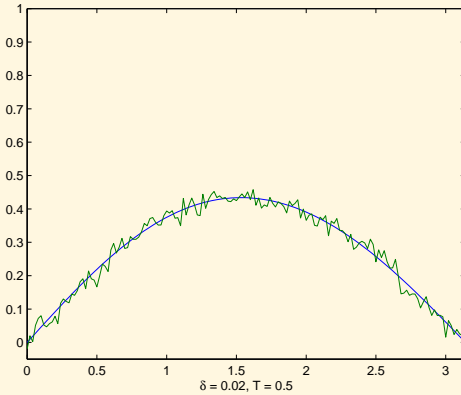
Sulje

Poistu



Sisältö:

1. Johdanto
2. Yleistetyt käännteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä



Kuva 1.9: a) Havainto $g(x, T = 0.50)$, b)-c) alkujakauman $f(x)$ rekonstruktio, kun summasta (1.15) on otettu vastaavasti mukaan $r = 2, 3$ ja 4 termiä. Katkiviivalla esitetty oikea ratkaisu. [\[KOODI\]](#)

Etusivu



Sivu 26 / 108

Takaisin

Full Screen

Sulje

Poistu



Ratkaisuhilan säätäminen ja katkaistun singulaariarvohajoitelman käyttäminen ovat kaksi perinteisintä regularisointimenetelmää. Nämä menetelmät ovat vielä laajalti käytössä, vaikka niiden ominaisuudet eivät yleensä ole kovin hyvät. Luvussa 3 esitettävä Tihonov-regularisointi on menetelmäluokka, joka on erittäin joustava ja sallii kohteen erityisominaisuuksien liittämisen estimointimenetelmään perinteisiä menetelmiä paremmin.

Tehtäviä

Sisältö:

1. *Johdanto*
2. *Yleistetyt käänteiskuvaukset*
3. *Tihonov-regularisointi*
4. *Iteratiiviset menetelmät*
5. *Tehtäviä*

Etusivu



Sivu 27 / 108

Takaisin

Full Screen

Sulje

Poistu



Yleistetyt käänteiskuvaukset

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 28 / 108

Takaisin

Full Screen

Sulje

Poistu

2.1. Linearisista operaattoreista

Operaattorilla T tarkoitetaan kuvausta, joka on määritelty jonkin lineaariavaruuden X osajoukossa (määrittelyjoukko $\mathcal{D}(T)$) ja joka saa arvonsa toisessa lineaariavaruudessa Y . Usein (varsinkin lineaarisilla T) $\mathcal{D}(T)$ on aliavaruus.

Esimerkki 2.1 *Olkoon derivaattaoperaattori D määritelty lineaariavaruuden $C([0, 1])$ osajoukossa $C^1([0, 1])$ siten, että $D : C^1([0, 1]) \mapsto C([0, 1])$, $Df = f'$. \square*

Nimityksistä huomautettakoon, että edellisessä esimerkissä operaattorin määrittelyjoukon alkiot ovat jatkuvasti derivoituvia funktioita, jotka on määritelty *alueessa* (välillä) $\Omega = [0, 1]$.

Olkoon $\mathcal{D}(T)$ operaattorin T määrittelyjoukko. Operaattori T on lineaarinen, jos

$$T(c_1 f_1 + c_2 f_2) = c_1 T f_1 + c_2 T f_2 \quad (2.1)$$



kaikille skalaareille c_1, c_2 ja vektoreille $f_1, f_2, c_1f_1 + c_2f_2 \in \mathcal{D}(T)$. Operaattori T on jatkuva pisteessä f , jos

$$\|Tf - Tf_k\| \rightarrow 0$$

kaikille jonoille $\{f_k\}$, joille $\|f - f_k\| \rightarrow 0$. Siten jatkuvuus riippuu normista. Lineaarioperaattori $T : \mathcal{D}(T) \mapsto Y$ on rajoitettu, jos on olemassa $M < \infty$ siten, että

$$\|Tf\|_Y \leq M\|f\|_X \quad \text{kaikille } f \in \mathcal{D}(T), \quad (2.2)$$

missä normit ovat vastaavien avaruuksien normeja. Rajoitetun kuvauksen normi on

$$\|T\| = \sup_{f \neq 0} \frac{\|Tf\|}{\|f\|} = \sup_{\|f\|=1} \|Tf\|. \quad (2.3)$$

Operaattorin $T : \mathcal{D}(T) \mapsto Y$ nolla-avaruus $\mathcal{N}(T)$ ja maalijoukko $\mathcal{R}(T)$ ovat joukot

$$\begin{aligned} \mathcal{N}(T) &= \{f \in \mathcal{D}(T) \subset X : Tf = 0\} \quad \text{ja} \\ \mathcal{R}(T) &= \{g \in Y : Tf = g \text{ jollekin } f \in \mathcal{D}(T)\}. \end{aligned}$$

Jatkossa oletetaan (ellei toisin mainita), että $\mathcal{D}(T) = X$.

Tässä luvussa tullaan huomaamaan, että käänteisongelmien ongelmana on, etteivät suorien ongelmien operaattoreiden T nolla-avaruudet ole "selviä". Toisin sanoen on olemassa vektoreita f siten, että $\|f\| = 1$, mutta $\|Tf\| \approx 0$, joten $f \notin \mathcal{N}(T)$.

Esimerkki 2.2 Olkoon $T : \mathbb{R}^N \mapsto \mathbb{R}^M$. Kuvausta vastaavaa matriisiä merkitään myös $T \in \mathbb{R}^{M \times N}$. Nyt

$$Tf = (v_1, v_2, \dots, v_N)(f_1, \dots, f_N)^T = \sum_{k=1}^N f_k v_k,$$

missä $v_k \in \mathbb{R}^M$ ovat matriisin T sarakkeet. Siten $\mathcal{R}(T) = \text{span}\{v_1, \dots, v_N\}$, mutta $\dim \mathcal{R}(T) \leq N$. Jos $\dim \mathcal{R}(T) < N$, on kuvauksella T ei-triviaali nolla-avaruus $\mathcal{N}(T)$. Mieti, miten nolla-avaruus ja maalijoukko liittyvät T :n sarakkeiden lineaariseen riippumattomuuteen ja kysymyksen, onko $M > N$?

Etusivu



Sivu 29 / 108

Takaisin

Full Screen

Sulje

Poistu

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä



Esimerkki 2.3 Olkoon $T \in \mathbb{R}^{2 \times 2}$ siten, että

$$T = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}.$$

Tällöin voidaan valita vektori $f = (-2, 1)^T$, jolloin $Tf = (0, 0)^T$, eli $f \in \mathcal{N}(T)$. Toisaalta myös $cf \in \mathcal{N}(T)$, $\forall c \in \mathbb{R}$. Helposti havaitaan, että $\dim \mathcal{R}(T) \neq 0$, joten $\dim \mathcal{N}(T) = 1$. Eli tässä tapauksessa $\mathcal{N}(T) = \text{span}\{(-2, 1)^T\}$.

Operaattori T on kompakti, jos jokaisen rajoitetun joukon $B \subset \mathcal{D}(T)$ kuvan sulkeuma $\overline{T(B)}$ on kompakti. Jos $\dim \mathcal{R}(T) < \infty$, on T kompakti. Siten periaatteessa kaikki kuvaukset $T \in \mathbb{R}^{M \times N}$ ovat kompakteja. Tähän seikkaan liittyviin ongelmiin palataan myöhemmin tässä luvussa.

2.1.1. Hilbert-avaruus

Hilbert-avaruudet \mathcal{H} ovat täydellisiä normiavaruuksia, joissa normi on määritelty sisätulon avulla siten, että

$$\|f\| = \langle f, f \rangle^{1/2}, \quad f \in \mathcal{H}.$$

Täydellisyydellä tarkoitetaan, että jokaisella suppenevalla Cauchy-jonolla $\{f_n\}$, jolle $\|f_n - f_m\| \rightarrow 0$, kun $m, n \rightarrow \infty$, $f_n \in \mathcal{H}$, on raja-arvo $f_n \rightarrow f \in \mathcal{H}$. Mielivaltaisen sisätuloavaruuden \mathcal{V} täydellistymä (sulkeuma) $\overline{\mathcal{V}}$ on kaikkien niiden jonojen $\{f_n\}$ raja-arvojen joukko, joille $\{f_n\} \subset \mathcal{V}$.

Esimerkki 2.4 Olkoon f_n välillä $[0, 2]$ määritelty funktio siten, että

$$f_n(x) = \begin{cases} 0 & \text{kun } x \in [0, 1 - 1/n] \\ nx & \text{kun } x \in [1 - 1/n, 1] \\ 1 & \text{kun } x \in [1, 2] \end{cases}$$

Nyt f_n kuuluu jatkuvien funktioiden joukkoon C . Kuitenkaan jonon $\{f_n\}$ rajaarvo

$$f = \begin{cases} 0 & \text{kun } x \in [0, 1] \\ 1 & \text{kun } x \in [1, 2] \end{cases}$$

ei ole jatkuva funktio. Joten jatkuvien funktioiden joukko C ei ole täydellinen.

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 30 / 108

Takaisin

Full Screen

Sulje

Poistu



Lisäksi määritellään W -normi

$$\|f\|_W = \langle f, Wf \rangle^{1/2}, \quad (2.4)$$

missä W on positiivisemidefiniitti lineaarioperaattori. Epäyhtälöt

$$\begin{aligned} |\langle v, w \rangle| &\leq \|v\| \|w\| && \text{Cauchy-Schwartz,} \\ \|v + w\| &\leq \|v\| + \|w\| && \text{kolmioepäyhtälö ja} \\ \|v + w\|^2 + \|v - w\|^2 &= 2\|v\|^2 + 2\|w\|^2 && \text{suunnikassääntö} \end{aligned}$$

ovat hyödyllisiä. Lineaarioperaattori $T : \mathcal{H} \mapsto \mathbb{C}$ on lineaarinen funktionaali. Lineaarinen funktionaali on rajoitettu, jos ja vain jos se on jatkuva. Olkoon $T : \mathcal{H} \mapsto \mathbb{C}$ rajoitettu lineaarinen funktionaali, missä \mathcal{H} on Hilbert-avaruus. Silloin Rieszin esitysauseen perusteella on olemassa yksikäsitteinen vektori $g \in \mathcal{H}$ siten, että

$$T(f) = \langle f, g \rangle \quad (2.5)$$

kaikille $f \in \mathcal{H}$. Lisäksi $\|T\| = \|g\|$.

Esimerkki 2.5 *Olkoon $T : L_1([0, 1]) \mapsto \mathbb{R} : Tf = \int_0^1 f(x) dx$. Osoita, että tämä on rajoitettu lineaarinen funktionaali. Mikä on sisätulo ja vektori $g \in L_1([0, 1])$? Oletetaan, että $f \in L_1$. Tällöin*

$$\int_0^1 |f(x)| dx = \|f\|_1 < \infty.$$

Sisätulo

$$T(f) = \langle f, g \rangle = \int_0^1 f(x)g(x)dx = \int_0^1 f(x)dx,$$

jos ja vain jos $g(x) \equiv 1$, kun $x \in [0, 1]$. \square

Olkoon $\mathcal{V} \subset \mathcal{H}$ ja $f \in \mathcal{H}$. Joukkoa $\mathcal{V}_f = \{w \mid w = v + f, v \in \mathcal{V}\}$ kutsutaan lineaariseksi monistoksi.

Sisältö:

1. Johdanto
2. Yleistetyt käännekuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 31 / 108

Takaisin

Full Screen

Sulje

Poistu



Olkoon $T : \mathcal{H}_1 \mapsto \mathcal{H}_2$ rajoitettu lineaarioperaattori, missä \mathcal{H}_1 ja \mathcal{H}_2 ovat Hilbert-avaruuksia. Operaattorin T adjungaatti T^* on operaattori $T^* : \mathcal{H}_2 \mapsto \mathcal{H}_1$, jolle pätee

$$\langle Tf, g \rangle_{\mathcal{H}_2} = \langle f, T^*g \rangle_{\mathcal{H}_1} \quad (2.6)$$

kaikille $f \in \mathcal{H}_1$ ja $g \in \mathcal{H}_2$. Jos $T = T^*$, on T itseadjungoitu (vastaava matriisi symmetrinen, jos avaruudet ja T reaalisia).

Esimerkki 2.6 Matriisin $T \in \mathbb{R}^{M \times N}$ adjungaatti. Olkoon $f \in \mathbb{R}^N$ ja $g, Tf \in \mathbb{R}^M$. Sisätulon ominaisuuksien perusteella

$$\begin{aligned} \langle Tf, g \rangle_{\mathbb{R}^M} &= \langle g, Tf \rangle_{\mathbb{R}^M} = g^T(Tf) \in \mathbb{R} \\ &= ((Tf)^T)^T = (f^T T^T)^T = (T^T)^T f \\ &= \langle T^T g, f \rangle_{\mathbb{R}^N} = \langle f, T^T g \rangle_{\mathbb{R}^N} \\ &= \langle f, T^*g \rangle_{\mathbb{R}^N} \quad \forall f, g, \end{aligned}$$

joten $T^* = T^T$. \square

Olkoon $\mathcal{V} \subset \mathcal{H}$ osajoukko. Joukon \mathcal{V} ortogonaalinen komplementti \mathcal{V}^\perp on niiden alkien $f \in \mathcal{H}$ joukko, joille pätee $\langle f, v \rangle = 0$ kaikille $v \in \mathcal{V}$.

Esimerkki 2.7 Olkoon $\mathcal{H} = \mathbb{R}^2$, $v = (1, 1)^T$ ja $\mathcal{V} = \text{span}\{v\}$. Tällöin, jos $x \in \mathcal{V}$, niin $x = cv$, $c \in \mathbb{R}$. Määrätään \mathcal{V}^\perp . Nyt $\langle f, v \rangle = f^T v = f_1 + f_2 = 0$, joten $f_1 = -f_2$. Siten $\mathcal{V}^\perp = \text{span}\{(1, -1)^T\}$, kuva 2.1. \square

2.1.2. Projektiot

Projektioperaattori (projektor) $P : \mathcal{H} \mapsto \mathcal{H}$ on operaattori, jolle pätee $P^2 = P$. Projektori on ortogonaaliprojektori, jos lisäksi $P = P^*$. Jollei toisin mainita, jatkossa oletetaan, että kaikki projektorit ovat ortogonaaliprojektoreita.

Sisältö:

1. Johdanto
2. Yleistetyt käännekuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 32 / 108

Takaisin

Full Screen

Sulje

Poistu



Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



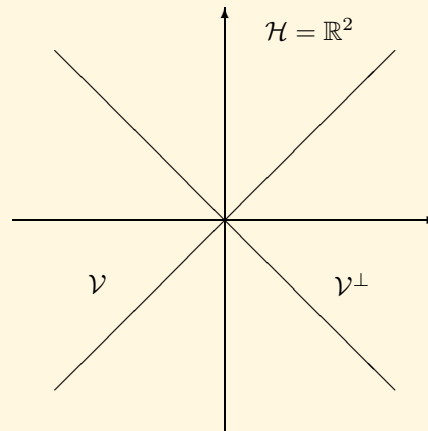
Sivu 33 / 108

Takaisin

Full Screen

Sulje

Poistu



Kuva 2.1: Ortogonaalinen komplementti.



Lause 2.1 Olkoon \mathcal{H} Hilbert-avaruus ja $\mathcal{V} \subset \mathcal{H}$ suljettu aliavaruus. Jokaiselle $f \in \mathcal{H}$ on olemassa yksikäsitteinen $v \in \mathcal{V}$ siten, että $\|f - v\| \leq \|f - w\|$ kaikille $w \in \mathcal{V}$. Tällöin $f - v \perp \mathcal{V}$, eli $f - v \in \mathcal{V}^\perp$. \square

Vektoria v kutsutaan vektorin f ortogonaaliseksi projektioksi aliavaruudelle \mathcal{V} ja merkitään $v = P_{\mathcal{V}}f$.

Esimerkki 2.8 Muodostetaan vektorin f projektiio $P_{\mathcal{V}}f$ normeeratulle vektorille v . Nyt

$$\langle f - P_{\mathcal{V}}f, v \rangle = \langle f - cv, v \rangle = 0, \quad c \in \mathbb{R}. \quad \text{Siten } \langle f, v \rangle = \langle cv, v \rangle = c \langle v, v \rangle = c \text{ ja } P_{\mathcal{V}}f = \langle f, v \rangle v. \quad \square$$

Esimerkki 2.9 Muodostetaan vektorin f projektiio $P_{\mathcal{V}}f$ ortonormaalille vektorijoukolle $\mathcal{V} = \text{span}\{v_k\}$. Koska $P_{\mathcal{V}}f \in \mathcal{V}$, voidaan kirjoittaa $P_{\mathcal{V}}f = \sum_k c_k v_k$. Vaatimuksen $f - P_{\mathcal{V}}f \perp \mathcal{V}$ perusteella asetetaan erityisesti $f - P_{\mathcal{V}}f \perp v_j$ kaikille j . Tämä on riittävää seuraavalla perusteella: Jos $f - P_{\mathcal{V}}f \perp v_j$ kaikille v_j , niin $f - P_{\mathcal{V}}f \perp \sum_j b_j v_j$ kaikille b_j . Mutta $\text{span}\{v_j\} = \mathcal{V}$, joten $f - P_{\mathcal{V}}f \perp \mathcal{V}$. Kuten edellisessä esimerkissä, saadaan siten

$$\begin{aligned} \langle f - P_{\mathcal{V}}f, v_j \rangle &= \left\langle f - \sum_k c_k v_k, v_j \right\rangle = 0, \quad \text{kaikille } j, \\ \langle f, v_j \rangle &= \sum_k c_k \underbrace{\langle v_k, v_j \rangle}_{\delta_{kj}} = c_j \quad \text{ja} \\ &= \begin{cases} 1, & k = j \\ 0, & k \neq j \end{cases} \\ P_{\mathcal{V}}f &= \sum_k c_k v_k \\ &= \sum_k \langle f, v_k \rangle v_k. \end{aligned}$$

Jos $\bar{\mathcal{V}} = \mathcal{H}$, on $P_{\mathcal{V}}f = f$ ja yllä olevaa muotoa kutsutaan vektorin f Fourier-sarjaksi kannan $\{v_j\}$ suhteen. \square

Etusivu



Sivu 34 / 108

Takaisin

Full Screen

Sulje

Poistu



Lause 2.2 Olkoon $T : \mathcal{H}_1 \mapsto \mathcal{H}_2$, missä \mathcal{H}_1 ja \mathcal{H}_2 ovat Hilbert-avaruuksia. Operaattorin T maalijoukolle ja nolla-avaruudelle pätee

$$\mathcal{N}(T^*) = \mathcal{R}(T)^\perp \quad \text{ja} \quad (2.7)$$

$$\mathcal{N}(T)^\perp = \mathcal{R}(T^*) . \quad (2.8)$$

TODISTUS. Olkoon $g^* \in \mathcal{N}(T^*) \subset \mathcal{H}_2$ ja $g \in \mathcal{R}(T)$. Silloin on olemassa $f \in \mathcal{H}_1$ siten, että $g = Tf$. Siten

$$\langle g, g^* \rangle = \langle Tf, g^* \rangle = \langle f, T^*g^* \rangle = 0 ,$$

sillä $g^* \in \mathcal{N}(T^*)$. Koska $g \in \mathcal{R}(T)$ ja $\langle g, g^* \rangle = 0$, niin $g^* \in \mathcal{R}(T)^\perp$ ja näin ollen

$$\mathcal{N}(T^*) \subset \mathcal{R}(T)^\perp .$$

Olkoon nyt $g^* \in \mathcal{R}(T)^\perp$. Silloin kaikille $f \in \mathcal{D}(T) \subset \mathcal{H}_1$ pätee

$$0 = \langle Tf, g^* \rangle = \langle f, T^*g^* \rangle$$

joten $T^*g^* = 0$ ja $g^* \in \mathcal{N}(T^*)$. Eli

$$\mathcal{R}(T)^\perp \subset \mathcal{N}(T^*) .$$

Siten

$$\mathcal{R}(T)^\perp = \mathcal{N}(T^*) .$$

Tulos (2.8) todistetaan vastaavasti ([harjoitustehtävä](#)). \square

Esimerkki 2.10 Tarkastellaan tapausta $T : \mathcal{H}_1 \mapsto \mathcal{H}_2$, $\mathcal{H}_1, \mathcal{H}_2 = \mathbb{R}^2$. Olkoon

$$T = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} = (v_1, v_2) ,$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 35 / 108

Takaisin

Full Screen

Sulje

Poistu



missä (v_1, v_2) on matriisi, jonka sarakkeet ovat v_1 ja v_2 . Määritään edellisessä lauseessa esiintyneet aliavaruudet. Olkoon $f \in \mathbb{R}^2$, jolloin $Tf = f_1v_1 + f_2v_2$. Siten

$$\mathcal{R}(T) = \{f_1v_1 + f_2v_2 \mid f = (f_1, f_2) \in \mathcal{D}(T) = \mathbb{R}^2\}.$$

Nyt $v_2 = 2v_1$, joten $Tf = cv_1$, $c \in \mathbb{R}$ ja

$$\mathcal{R}(T) = \text{span}\{v_1, v_2\} = \text{span}\{v_1\} = \{c(1, 1), c \in \mathbb{R}\}.$$

Olkoon nyt $f \in \mathcal{N}(T)$, joten $Tf = 0$. Edellisen perusteella saadaan

$$f_1v_1 + f_2v_2 = f_1v_1 + 2f_2v_1 = (f_1 + 2f_2)v_1 = 0.$$

Siten $f_1 = -2f_2$ ja

$$\mathcal{N}(T) = \{c(-2, 1), c \in \mathbb{R}\}.$$

Vastaavasti adjungaatile $T^* = T^T$ saadaan, että

$$\mathcal{R}(T^*) = \{c(1, 2), c \in \mathbb{R}\}$$

ja

$$\mathcal{N}(T^*) = \{c(1, -1), c \in \mathbb{R}\},$$

kuva 2.2. \square

Esimerkki 2.11 Tarkastellaan Gram-Schmidt -ortogonalisointimenetelmää. Olkoon

$$\mathcal{V} = \text{span}\{v_k, k = 1, \dots, r\} \subset \mathcal{H}, \quad \dim \mathcal{H} \geq r$$

linearisesti riippumaton vektorijoukko. Tuotetaan tästä joukosta ortonormaali kanta aliavaruudelle \mathcal{V} rekursiivisesti. Oletetaan, että \mathcal{V} on vähintään kaksiulotteinen sisätuloavaruus ja joukko $\{v_1, \dots, v_r\}$ sen jokin kanta. Ortogonaalisointi tapahtuu seuraavasti:

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 36 / 108

Takaisin

Full Screen

Sulje

Poistu



Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



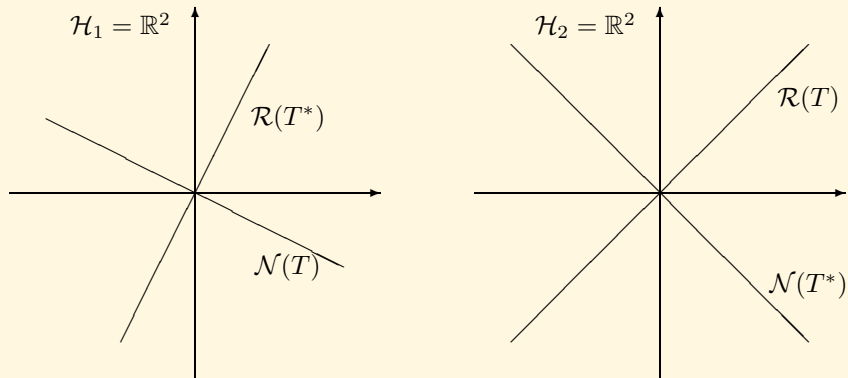
Sivu 37 / 108

Takaisin

Full Screen

Sulje

Poistu



Kuva 2.2: Matriisin T maalijoukko ja nolla-avaruus.



- Korvataan v_1 yksikkövektorilla

$$u_1 := \frac{v_1}{\|v_1\|}.$$

- Korvataan v_2 sen normitetulla ortogonaaliprojektiolla

$$\tilde{u}_2 := v_2 - \langle v_2, u_1 \rangle u_1,$$

jolloin

$$u_2 := \frac{\tilde{u}_2}{\|\tilde{u}_2\|}.$$

- Toistetaan arvoilla $k = 1, \dots, r$. Korvataan v_k joka kierroksella sen normitetulla ortogonaaliprojektiolla

$$\tilde{u}_k := v_k - \sum_{i=1}^{k-1} \langle v_k, u_i \rangle u_i,$$

jolloin

$$u_k := \frac{\tilde{u}_k}{\|\tilde{u}_k\|}.$$

2.1.3. Klassinen ja minimiresiduaaliratkaisu

Lause 2.3 Olkoon $\mathcal{V} \subset \mathcal{H}$ suljettu aliavaruus. Silloin $\mathcal{H} = \mathcal{V} \oplus \mathcal{V}^\perp$ (suora summa), missä hajotelma on yksikäsitteinen. \square

Lauseen merkitys on siinä, että jokainen $f \in \mathcal{H}$ voidaan esittää muodossa $f = f_1 + f_2$, $f_1 \in \mathcal{V}$, $f_2 \in \mathcal{V}^\perp$ yksikäsitteisesti. Olkoon $f \in \mathcal{H}$. Nyt voidaan kirjoittaa

$$f = If = (P_{\mathcal{V}} + (I - P_{\mathcal{V}}))f = \underbrace{P_{\mathcal{V}}f}_{\in \mathcal{V}} + \underbrace{(I - P_{\mathcal{V}})f}_{\in \mathcal{V}^\perp},$$

missä $I : \mathcal{H} \mapsto \mathcal{H}$ on identiteettioperaattori.

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 38 / 108

Takaisin

Full Screen

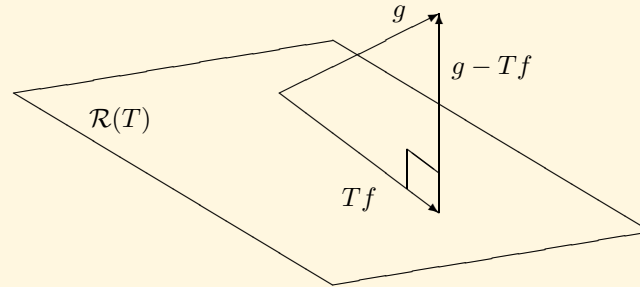
Sulje

Poistu



Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä



Kuva 2.3: LS-ratkaisu: datan projektio operaattorin T maalijoukolle.

Lause 2.4 Olkoon $T : \mathcal{H}_1 \mapsto \mathcal{H}_2$, $f \in \mathcal{D}(T)$ ja $g \in \mathcal{H}_2$. Tällöin residuaali $\|Tf - g\|$ saavuttaa miniminsä, kun $Tf - g \in \mathcal{R}(T)^\perp$, kuva 2.3. Todistus on harjoitustehtävä. \square

Koska $\mathcal{R}(T)^\perp = \mathcal{N}(T^*)$, saadaan edellisen lauseen minimille karakterisointi, jota kutsutaan *normaaliyhtälöiksi*

$$T^*(Tf - g) = 0 \Leftrightarrow \quad (2.9)$$

$$T^*Tf = T^*g. \quad (2.10)$$

Tämän yhtälön ratkaisua kutsutaan myös LS-ratkaisuksi (*Least Squares*).

Jos $\mathcal{N}(T) = \{0\}$, niin $\mathcal{N}(T^*T) = \{0\}$ ($\mathcal{N}(T) = \mathcal{N}(T^*T)$, [laskuharjoitukset](#)) ja voidaan kirjoittaa

$$f = (T^*T)^{-1}T^*g. \quad (2.11)$$

Jos lisäksi $g \in \mathcal{R}(T)$, on olemassa $v \in \mathcal{D}(T)$ siten, että $Tv = g$, joten $\|Tv - g\| = 0$. Tällöin myös

$$T^*Tf = T^*g = T^*Tv,$$

Etusivu



Sivu 39 / 108

Takaisin

Full Screen

Sulje

Poistu



joten $f = v$. Toisaalta, jos $\mathcal{R}(T) \neq \mathcal{H}_2$ ja/tai $\mathcal{N}(T) \neq \{0\}$, ei T ole bijektio eikä (tavallista) käänteiskuvausta voida määritellä.

Olkoon nyt kuitenkin $g \in \mathcal{R}(T)$, jolloin $Tv = g$. Vektori $v \in \mathcal{H}_1$ on siten yhtälön $Tv = g$ klassinen ratkaisu. Jos $\mathcal{N}(T) = \{0\}$, on v yksikäsitteinen. Jos $\mathcal{N}(T) \neq \{0\}$, kirjoitetaan $v = v_1 + v_0$, missä $v_1 \in \mathcal{N}(T)^\perp$ ja $v_0 \in \mathcal{N}(T)$. Nyt

$$Tv = T(v_1 + v_0) = Tv_1 + Tv_0 = Tv_1 = g,$$

joten v ei ole yksikäsitteinen. Jos v on jokin yhtälön $Tv = g$ ratkaisu, on siten myös $v + v_0$ tämän yhtälön ratkaisu, missä v_0 on mielivaltainen (mielivaltaisen suuri!) operaattorin T nollavaruuden alkio. Nyt

$$\|v\|^2 = \|v_1\|^2 + \|v_0\|^2 \geq \|v_1\|^2,$$

joten kaikista ratkaisuista v on normiltaan pienin se, jolle $v_0 = 0$. Kuitenkaan operaattorin T^*T käänteisoperaattoria ei ole olemassa, joten ratkaisua v ei saada normaaliyhtälöiden ratkaistun muodon (2.11) avulla. Siten tehtävän ”Määrää f , kun $Tf = g$ ja g tunnettu.” ratkaisu ja tulkinta riippuu seuraavista kysymyksistä

$$\begin{aligned} \mathcal{R}(T) &\stackrel{?}{=} \mathcal{H}_2, \\ \mathcal{N}(T) &\stackrel{?}{=} \{0\} \quad \text{ja} \\ g &\stackrel{?}{\in} \mathcal{R}(T). \end{aligned}$$

Seuraavissa kappaleissa tarkastellaan, mitä käsitteellä *yleistetty ratkaisu* tarkoitetaan.

2.2. Singulaariarvohajotelma

Tarkastellaan yhtälöä

$$Kf = g, \tag{2.12}$$

missä $K : \mathcal{H}_1 \mapsto \mathcal{H}_2$ on kompakti lineaarioperaattori. Pyritään esittämään f ja g operaattoriin K liittyvien (yleistettyjen) Fourier-sarjojen avulla. Avaruus \mathcal{H} on separoituva, jos se sisältää

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 40 / 108

Takaisin

Full Screen

Sulje

Poistu



numeroituvan täydellisen ortonormaalien joukon. Oletetaan, että \mathcal{H}_1 ja \mathcal{H}_2 ovat separoituvia, jolloin alkioihin f ja g näiden avaruuksien täydellisten ortonormaalien kantojen suhteen kehitetyt Fourier-sarjat suppenevat normien mielessä.

Operaattorit $K^*K : \mathcal{H}_1 \mapsto \mathcal{H}_1$ ja $KK^* : \mathcal{H}_2 \mapsto \mathcal{H}_2$ ovat selvästi kompakteja, itseadjungoituvia (symmetrisiä) ja niiden nollassa eroavat ominaisarvot λ_k ovat samoja ja positiivisia. Tämä nähdään seuraavasti

$$\begin{aligned} K^*Kv &= \lambda v , \\ KK^*(Kv) &= \lambda(Kv) \quad \text{ja} \\ KK^*w &= \lambda w , \end{aligned}$$

missä $w = Kv$. Olkoot $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$ nämä ominaisarvot (viivaspektri) ja $\{v_k\}$ ominaisarvoja vastaavien operaattorin K^*K ortonormaalien ominaisvektorien joukko.. Joukko $\{v_k\}$ on täydellinen aliavaruudessa $\mathcal{R}(K^*K) = \mathcal{N}(K^*K)^\perp = \mathcal{N}(K)^\perp \subset \mathcal{H}_1$. Olkoon $\sigma_k = \sqrt{\lambda_k}$ ja $u_k = \sigma_k^{-1}Kv_k$. Siten

$$K^*u_k = \sigma_k v_k , \quad (2.13)$$

$$Kv_k = \sigma_k u_k \quad \text{ja} \quad (2.14)$$

$$K^*Kv_k = \sigma_k K^*u_k = \sigma_k^2 v_k = \lambda_k v_k . \quad (2.15)$$

Vastaavasti operaattorin KK^* ortonormaalien ominaisvektoreiden joukko $\{u_k\}$ on täydellinen aliavaruudessa $\mathcal{R}(KK^*) = \mathcal{N}(KK^*)^\perp = \mathcal{N}(K^*)^\perp \subset \mathcal{H}_2$.

Systeemiä $\{\sigma_k, v_k, u_k\}$ kutsutaan operaattorin K *singulaarisysteemiksi* ja K :n jatkossa esitettävää näihin suureisiin perustuvaa hajotelmaa *singulaariarvohajotelmaksi*. Luvut σ_k ovat K :n *singulaariarvot*.

Vektori $f \in \mathcal{H}_1$ voidaan nyt esittää muodossa

$$f = \underbrace{P_{\mathcal{N}(K)}f}_{\in \mathcal{N}(K)} + \underbrace{\sum_{k=1}^{\infty} \langle f, v_k \rangle v_k}_{\in \mathcal{N}(K)^\perp} , \quad (2.16)$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 41 / 108

Takaisin

Full Screen

Sulje

Poistu



Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



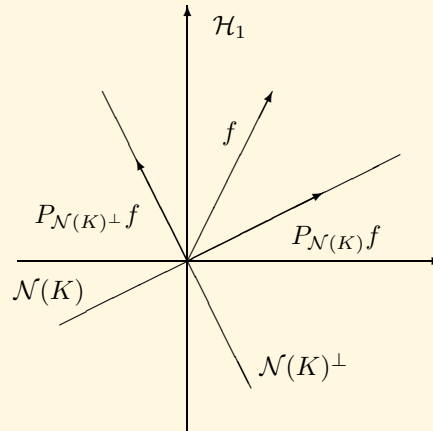
Sivu 42 / 108

Takaisin

Full Screen

Sulje

Poistu



Kuva 2.4: Vektorin f projektiot aliavaruuksille $\mathcal{N}(K)$ ja $\mathcal{N}(K)^\perp$.

missä $P_{\mathcal{A}}$ on ortogonaaliprojektori avaruudelle \mathcal{A} , kuva 2.4. Siten

$$Kf = \underbrace{KP_{\mathcal{N}(K)}f}_{=0} + \sum_{k=1}^{\infty} \langle f, v_k \rangle Kv_k \quad (2.17)$$

$$= \sum_{k=1}^{\infty} \sigma_k \langle f, v_k \rangle u_k, \quad (2.18)$$

jota kutsutaan operaattorin K *singulaariarvohajotelmaksi (SVD)*.



2.2.1. Klassinen ja miniminormiratkaisu

Mikäli yhtälöllä (2.12) on klassinen ratkaisu f , on oltava $g \in \mathcal{R}(K)$. Tällöin

$$\langle g, u_k \rangle = \langle Kf, \sigma_k^{-1} K v_k \rangle = \langle f, \sigma_k^{-1} K^* K v_k \rangle = \sigma_k \langle f, v_k \rangle, \quad (2.19)$$

koska $K^* K v_k = \lambda_k v_k = \sigma_k^2 v_k$. Edelleen¹

$$\sum_{k=1}^{\infty} \lambda_k^{-1} |\langle g, u_k \rangle|^2 = \sum_{k=1}^{\infty} |\langle f, v_k \rangle|^2 \leq \|f\|^2 < \infty. \quad (2.20)$$

Siten erääksi ratkaisuksi saadaan

$$\begin{aligned} f &= \sum_{k=1}^{\infty} \langle f, v_k \rangle v_k \\ &= \sum_{k=1}^{\infty} \sigma_k^{-1} \langle g, u_k \rangle v_k. \end{aligned}$$

Jos $\mathcal{N}(K) \neq \{0\}$, ovat kaikki muotoa

$$f = \underbrace{\sum_{k=1}^{\infty} \frac{\langle g, u_k \rangle}{\sigma_k} v_k}_{\in \mathcal{N}(K)^\perp} + f_0, \quad (2.21)$$

missä $f_0 \in \mathcal{N}(K)$, ovat vektorit ovat yhtälön (2.12) ratkaisuja. Koska

$$\sum_{k=1}^{\infty} \frac{\langle g, u_k \rangle}{\sigma_k} v_k \perp f_0, \quad (2.22)$$

¹Besselin epäyhtälö: $\sum_{k=1}^{\infty} |\langle f, v_k \rangle|^2 \leq \|f\|^2$ ortonormaalille joukolle $\{v_k\}$. Yhtäsuuruus toteutuu, kun $\text{span}\{v_k\} = \mathcal{H}$, mutta yleisesti $\text{span}\{v_k\} \subset \mathcal{H} \ni f$.

Etusivu



Sivu 43 / 108

Takaisin

Full Screen

Sulje

Poistu

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä



on yhtälön (2.12) ratkaisuista normiltaan pienin se, jolle $f_0 = 0$. Tätä ratkaisua kutsutaan *miniminnormiratkaisuksi*. Siten, jos $\mathcal{N}(K)^\perp \neq \mathcal{H}_1$, eli $\mathcal{N}(K) \neq \{0\}$, on yhtälön (2.12) toteuttavia ratkaisuja äärettömän monta, kuva 2.5.

Ongelmalla voi siis olla klassinen ratkaisu, mutta se ei ole välttämättä yksikäsitteinen. Toisaalta kappaleessa 2.1.3 käsiteltiin tapausta, jossa klassista ratkaisua ei välttämättä ollut. Miten käsitellään tapaus, jossa klassista ratkaisua ei ole eikä yleistettykään ratkaisu (LS) ole yksikäsitteinen? Tähän kysymykseen vastaa pseudoinvertssikuvausten teoria.

2.3. Pseudoinverssi

Tarkastellaan edelleen yhtälön

$$Kf = g, \quad K : \mathcal{H}_1 \mapsto \mathcal{H}_2 \quad (2.23)$$

ratkaisujen olemassaoloa ja laatua. Maaliavaruuden määritelmän perusteella on selvää, että f voi olla yhtälön *klassinen* ratkaisu vain, jos $g \in \mathcal{R}(K)$. Useissa tilanteissa kuitenkin $\mathcal{R}(K) \neq \mathcal{H}_2$, jolloin klassista ratkaisua ei yleensä ole olemassa. Tällainen tilanne on esimerkiksi äärellisdimensioisen ylideterministisen yhtälöryhmän ratkaiseminen. Vektoria $f \in \mathcal{H}_1$ kutsutaan yhtälön (2.23) *LS*-ratkaisuksi, jos

$$\|Kf - g\| = \min\{\|Ku - g\|, u \in \mathcal{H}_1\}. \quad (2.24)$$

Äärellisdimensioisessa tapauksessa tätä ratkaisua kutsutaan *pienimmän neliösomman* ratkaisuksi. Nyt $g \in \mathcal{H}_2 = \mathcal{R}(K) \oplus \mathcal{R}(K)^\perp$ ja voidaan kirjoittaa

$$g = P_{\mathcal{R}(K)}g + P_{\mathcal{R}(K)^\perp}g = P_{\mathcal{R}(K)}g + (I - P_{\mathcal{R}(K)})g$$

ja edelleen

$$Kf - g \in \mathcal{R}(K)^\perp = \mathcal{N}(K^*) \quad \text{ja} \quad (2.25)$$

$$K^*Kf = K^*g. \quad (2.26)$$

LS-ratkaisu on yksikäsitteinen, jos ja vain jos $\mathcal{N}(K) = \mathcal{N}(K^*K) = \{0\}$. Mitä tämä tarkoittaa, kun $K \in \mathbb{R}^{M \times N}$, $M > N$?

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 44 / 108

Takaisin

Full Screen

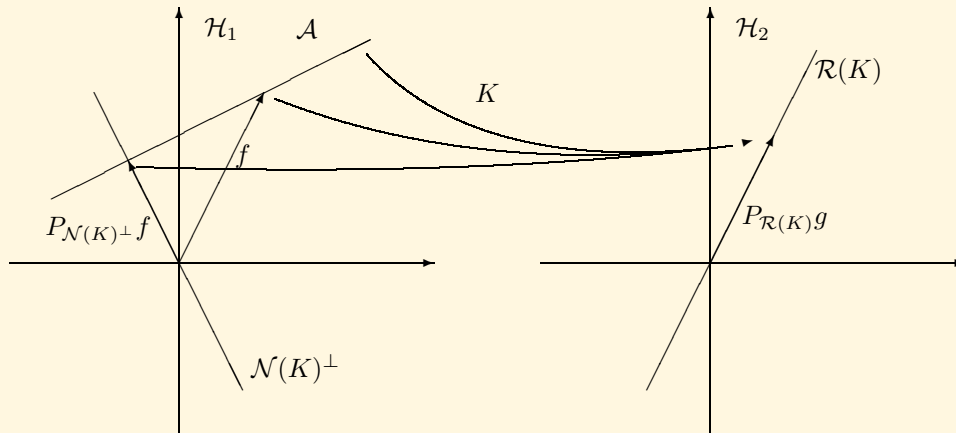
Sulje

Poistu



Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä



Kuva 2.5: Kaikkien LS-ratkaisujen muodostama monisto \mathcal{A} kuvautuu pisteeseen $P_{\mathcal{R}(K)}g$. Moniston on muotoa $\mathcal{A} = \{f \mid Kf = P_{\mathcal{R}(K)}g\} = P_{\mathcal{N}(K)^\perp}f + \mathcal{N}(K)$.

Etusivu



Sivu 45 / 108

Takaisin

Full Screen

Sulje

Poistu



2.3.1. Pseudoinverssin määritelmä ja karakterisointi

Kuvausta K^\dagger , joka liittää annetun $g \in \mathcal{D}(K^\dagger) = \mathcal{R}(K) \oplus \mathcal{R}(K)^\perp$ yksikäsitteiseen miniminormiseen LS-ratkaisuun, kutsutaan *yleistetyksi (Moore-Penrose) käänteiskuvaukseksi* tai *pseudoinverssiksi*. Siten²

$$K^\dagger g = \operatorname{sol} \min_u \{ \|u\| \mid u = \operatorname{sol} \min \|Ku - g\|, \quad u \in \mathcal{H}_1 \} . \quad (2.27)$$

Pseudoinverssille on olemassa useita ekvivalentteja määritelmiä, joista seuraavassa tärkeimmät.

- (Funktionaalinen määritelmä) Olkoon $K : \mathcal{H}_1 \mapsto \mathcal{H}_2$ ja $g \in \mathcal{H}_2$. Kuvauksen K pseudoinverssille K^\dagger pätee

$$\begin{aligned} K^\dagger g &= 0 & , g \in \mathcal{R}(K)^\perp & \text{ ja} \\ K^\dagger g &= (K|_{\mathcal{N}(K)^\perp})^{-1}g & , g \in \mathcal{R}(K) . \end{aligned}$$

Merkintä $K|_{\mathcal{N}(K)^\perp}$ tarkoittaa operaattorin K rajoittumaa aliavaruudelle $\mathcal{N}(K)^\perp$, missä termi *rajoittuma* viittaa määrittelyjoukkoon siten, että $\mathcal{D}(K|_{\mathcal{N}(K)^\perp}) = \mathcal{N}(K)^\perp \cap \mathcal{D}(K)$. Nyt siis $g \in \mathcal{D}(K^\dagger) = \mathcal{H}_2$ ja jaetaan g siten, että $g = g_1 + g_0$, missä $g_1 \in \mathcal{R}(K)$ ja $g_0 \in \mathcal{R}(K)^\perp$. Siten $K^\dagger g = K^\dagger g_1 + K^\dagger g_0 = K^\dagger g_1$, eli K^\dagger projisoi g :n “aluksi” maalijoukolle $\mathcal{R}(K)$ ja määrää sitten ongelman $Kf = P_{\mathcal{R}(K)}g \in \mathcal{R}(K)$ ratkaisun avaruudessa $\mathcal{N}(K)^\perp$.

- (Moore) Olkoon $K : \mathcal{H}_1 \mapsto \mathcal{H}_2$ ja $g \in \mathcal{H}_2$. Kuvauksen K pseudoinverssille K^\dagger pätee

$$\begin{aligned} KK^\dagger &= P_{\mathcal{R}(K)} \quad \text{ja} \\ K^\dagger K &= P_{\mathcal{R}(K^\dagger)} = P_{\mathcal{R}(K^*)} . \end{aligned}$$

- (Penrose) Olkoon $K : \mathcal{H}_1 \mapsto \mathcal{H}_2$ ja $g \in \mathcal{H}_2$. Kuvauksen K pseudoinverssille K^\dagger pätee

$$KK^\dagger K = K ,$$

²Merkintä sol luetaan “*solution of the (minimization) problem*”.

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 46 / 108

Takaisin

Full Screen

Sulje

Poistu



$$\begin{aligned}K^\dagger K K^\dagger &= K^\dagger, \\(K K^\dagger)^* &= K K^\dagger \text{ ja} \\(K^\dagger K)^* &= K^\dagger K.\end{aligned}$$

Esimerkki 2.12 (Funktionaalinen määritelmä) Oletetaan esimerkiksi, että

$$K = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}.$$

Tällöin $\mathcal{N}(K) = \text{span}\{(-2, 1)^\text{T}\}$ ja $\mathcal{N}(K)^\perp = \text{span}\{(1, 2)^\text{T}\}$. Oletetaan, että $f \in \mathcal{N}(K)^\perp$ jolloin f on muotoa

$$f = c \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

ja edelleen $K|_{\mathcal{N}(K)^\perp}$ -lle

$$K(cf) = c \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = c \begin{pmatrix} 5 \\ 5 \end{pmatrix}.$$

Voidaan osoittaa, että $\mathcal{R}(K) = \text{span}\{(1, 1)^\text{T}\}$. Merkitään $u = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^\text{T} \in \mathcal{R}(K)$ ja olkoon $g = (2, 1)^\text{T}$. Tällöin

$$P_{\mathcal{R}(K)}g = \langle g, u \rangle u = \frac{3}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Edelleen

$$\begin{aligned}K|_{\mathcal{N}(K)^\perp}(cf) = P_{\mathcal{R}(K)}g &\Leftrightarrow c \begin{pmatrix} 5 \\ 5 \end{pmatrix} = \frac{3}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &\Leftrightarrow c = \frac{3}{10}.\end{aligned}$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 47 / 108

Takaisin

Full Screen

Sulje

Poistu



Siis

$$K^\dagger g = K^\dagger \begin{pmatrix} 2 \\ 1 \end{pmatrix} = c \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{3}{10} \\ \frac{3}{5} \end{pmatrix}.$$

Pseudoinverssin määritelmä kattaa “kaikki tilanteet”:

- $\|Kf - g\| = 0$, eli $g \in \mathcal{R}(K)$ ja ratkaisu on klassinen.
 - Jos $\mathcal{N}(K) = \{0\}$, on $f = K^{-1}g = K^\dagger g$.
 - Jos $\mathcal{N}(K) \neq \{0\}$, on muotoa $f = f_1 + f_0$ olevia ratkaisuja ääretön määrä, missä $f_1 \in \mathcal{N}(K)^\perp$ ja $f_0 \in \mathcal{N}(K)$. Tällöin $K^\dagger g = f_1$ ja $Kf_1 = g$.
- $\min \|Kf - g\| \neq 0$, eli $g \notin \mathcal{R}(K)$ ja klassista ratkaisua ei ole olemassa. Nyt normin $\|Kf - g\| \neq 0$ minimoiville ratkaisuille $f = f_1 + f_0$ pätee $K^*Kf = K^*K(f_1 + f_0) = K^*g$, missä $f_1 \in \mathcal{N}(K)^\perp$ ja $f_0 \in \mathcal{N}(K)$.
 - $\mathcal{N}(K) = \mathcal{N}(K^*K) = \{0\}$, jolloin $f_0 = 0$ ja $f = K^\dagger g = (K^*K)^{-1}K^*g$ on yksikäsitteinen.
 - $\mathcal{N}(K) = \mathcal{N}(K^*K) \neq \{0\}$, jolloin kaikki muotoa $f = f_1 + f_0$ olevat vektorit ovat yhtälön $K^*K(f_1 + f_0) = K^*g$ ratkaisuja. Näistä on normiltaan pienin $K^\dagger g = f_1$.

Näin ollen käänteisongelmia karakterisoiva piirre on, että asiaankuuluvat nolla-avaruudet ja maalijoukot eivät ole “selviä”. Voi olla esimerkiksi niin, että eksaktisti tarkasteltuna $\mathcal{N}(K) = \{0\}$, mutta toisaalta on olemassa vektoreita, joille $\|v\| = 1$ ja $Kv \approx 0$.

Karkeasti ottaen voidaankin sanoa, että käänteisongelmat ja regularisointiteoria käsittelevät tapauksia, joissa miniminormiratkaisulla on vähän tai ei mitään tekemistä haettujen ratkaisujen kanssa.

Etusivu



Sivu 48 / 108

Takaisin

Full Screen

Sulje

Poistu



2.3.2. Minimimirroratkaisujen määrääminen “geneerisesti”

Tarkastellaan minimimirroratkaisun määräämistä, kun “käytettävissä” on ainoastaan ohjelma, jonka avulla voidaan laskea operaattorin maalijoukolle ortonormaali kanta. Matriisien tapauksessa tämä vastaa matriisin sarakkeiden ortonormeerausta esimerkiksi Gram-Schmidt -ortogonalisointimenetelmällä, jota voidaan toki käyttää muidenkin vektoriavaruuksien kuin \mathbb{R}^n :n tapauksessa.

Kuten edellä, ongelma on siis määrätä ongelman $\min \|Kf - g\|$ ratkaisuista f normiltaan pienin. Merkitään kuten edellä

$$g = \underbrace{g_1}_{\in \mathcal{R}(K)} + \underbrace{g_0}_{\in \mathcal{R}(K)^\perp} .$$

Residuaali on minimissään, kun $Kf = g_1 = P_{\mathcal{R}(K)}g$, kuva 2.6.

Miten g_1 saadaan?

1. Määrätään “efektiivinen” ortonormaali kanta $\{u_k\}$ maalijoukolle $\mathcal{R}(K)$ (Gram-Schmidt).

2. $g_1 = \sum_k \langle g, u_k \rangle u_k = P_{\mathcal{R}(K)}g$.

Seuraavaksi hajotetaan ratkaisut $\{f \mid Kf = g_1\}$

$$f = \underbrace{f_1}_{\in \mathcal{N}(K)^\perp} + \underbrace{f_0}_{\in \mathcal{N}(K)} ,$$

missä f_1 on yksikäsitteinen. Vektori f_0 on yksikäsitteinen vain, jos $\mathcal{N}(K) = \{0\}$, jolloin siis $f_0 = 0$. Koska ortogonaalisuuden perusteella $\|f\| = \|f_1\| + \|f_0\|$, on näistä ratkaisuista normiltaan pienin $f = f_1$, kuva 2.7.

Miten f_1 saadaan? Määrätään “efektiivinen” ortonormaali kanta $\{v_k\}$ aliavaruudelle $\mathcal{N}(K)^\perp = \mathcal{R}(K^*)$ (Gram-Schmidt). Nyt f_1 on muotoa $f_1 = \sum_k c_k v_k$ ja voidaan kirjoittaa

$$Kf_1 = \sum_k c_k \underbrace{Kv_k}_{\doteq w_k} = \sum_k c_k w_k = g_1 .$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 49 / 108

Takaisin

Full Screen

Sulje

Poistu



Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



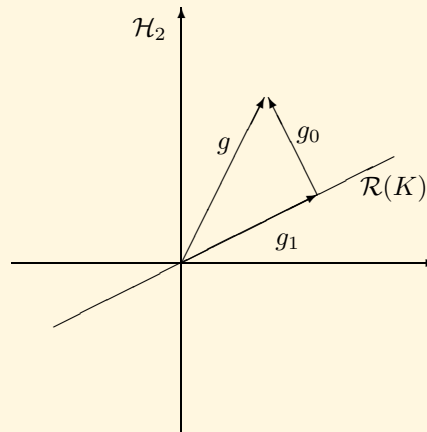
Sivu 50 / 108

Takaisin

Full Screen

Sulje

Poistu



Kuva 2.6: Datan g ortogonaali hajotelma avaruudessa \mathcal{H}_2 .



Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



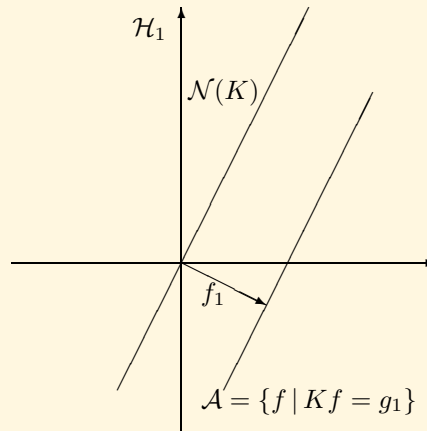
Sivu 51 / 108

Takaisin

Full Screen

Sulje

Poistu



Kuva 2.7: LS-ratkaisujen f ortogonaalinen hajotelma avaruudessa \mathcal{H}_1 .



Otetaan sisätulot viimeisimmästä yhtälöstä vektoreiden $w_j = Kv_j$ kanssa, jolloin saadaan yhtälöryhmä

$$\sum_k c_k \langle w_k, w_j \rangle = \langle g_1, w_j \rangle, \quad \text{kaikille } j.$$

Tämän (mahdollisesti äärettömän suuren) yhtälöryhmän ratkaisusta saadaan kertoimet c_k ja siten edelleen f_1 .

Jos joukko $\{v_k\}$ ja siten $\{w_k\}$ ovat äärellisiä, saadaan kertoimet c_k seuraavan yhtälöryhmän ratkaisusta

$$\begin{pmatrix} \langle w_1, w_1 \rangle & \cdots & \langle w_1, w_r \rangle \\ \vdots & \ddots & \vdots \\ \langle w_r, w_1 \rangle & \cdots & \langle w_r, w_r \rangle \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_r \end{pmatrix} = \begin{pmatrix} \langle g_1, w_1 \rangle \\ \vdots \\ \langle g_1, w_r \rangle \end{pmatrix}.$$

Merkitään tätä yhtälöryhmää $\Gamma c = \gamma$.

Nyt jos joukko $\{w_k\} \subset \mathcal{H}_2$ "sattuu" olemaan ortonormaali, saadaan kertoimet $c_j = \langle g_1, w_j \rangle$ suoraan. Edelleen, jos joukot $\{u_k\}$ ja $\{v_k\}$ olisi muodostettu erikoisesti siten, että $w_k = Kv_k = \sigma_k u_k$, saataisiin suoraan

$$\begin{aligned} g_1 &= \sum_k \langle g, u_k \rangle u_k \\ &= \sum_k \langle g, w_k \rangle w_k = \sum_k c_k \sigma_k u_k, \\ c_k &= \frac{\langle g, u_k \rangle}{\sigma_k} \quad \text{ja} \\ f_1 &= \sum_k \frac{\langle g, u_k \rangle}{\sigma_k} v_k, \end{aligned}$$

mikä on juuri edellä esitetty SVD-ratkaisu.

Pragmaattinen valinta avaruuden $\mathcal{N}(K)^\perp$ dimensioksi johtaa myös regularisointimenetelmään

Etusivu



Sivu 52 / 108

Takaisin

Full Screen

Sulje

Poistu



kuten kappaleessa 1.4. Tällöin asetetaan $\sigma_{r+1} = \sigma_{r+2} = \dots = 0$ ja

$$f_1 = \sum_{k=1}^r \frac{\langle g, u_k \rangle}{\sigma_k} v_k .$$

Menetelmää kutsutaan katkaistuun singulaariarvoajoitelmaan perustuvaksi regularisoinniksi. Ohessa **MATLAB-esimerkki** katkaisuun perustuvasta regularisoinnista.

2.3.3. Äärellisdimensioiset kuvaukset

Olkoon $K \in \mathbb{R}^{N \times M}$. Nyt voidaan kirjoittaa

$$K = USV^T, \quad (2.28)$$

missä $S = \text{diag}(\sigma_1, \dots, \sigma_r, 0, \dots, 0) \in \mathbb{R}^{N \times M}$ ja $r \leq \min\{M, N\}$ on matriisin K ranki. Lisäksi matriisien $V \in \mathbb{R}^{M \times M}$ ja $U \in \mathbb{R}^{N \times N}$ sarakkeet v_k ja u_k ovat ortonormaalit. Matriisin V sarakkeet virittävät avaruuden \mathbb{R}^M siten, että $\text{span}\{v_1, \dots, v_r\} = \mathcal{N}(K)^\perp$ ja $\text{span}\{v_{r+1}, \dots, v_M\} = \mathcal{N}(K)$. Vastaavasti matriisin U sarakkeet virittävät avaruuden \mathbb{R}^N ja $\text{span}\{u_1, \dots, u_r\} = \mathcal{R}(K)$ sekä $\text{span}\{u_{r+1}, \dots, u_N\} = \mathcal{R}(K)^\perp$.

Numeerisesti kannat operaattorin K nolla-avaruudelle ja maalijoukolle saadaan esimerkiksi ortogonalisointimenettelyjen avulla (Gram-Schmidt) ja käyttämällä hyväksi relaatioita (2.7–2.8), kuten edellisessä kappaleessa esitettiin.

Muodostakoon matriisin U_1 sarakkeet (matriisin U ensimmäiset r saraketta) ortonormaalin kannan maalijoukolle $\mathcal{R}(K)$. Nyt saadaan helposti (laskuharjoitukset), että

$$P_{\mathcal{R}(K)} = U_1(U_1^T U_1)^{-1} U_1^T = U_1 U_1^T \quad \text{ja} \quad (2.29)$$

$$P_{\mathcal{R}(K)^\perp} = I - U_1(U_1^T U_1)^{-1} U_1^T = I - U_1 U_1^T . \quad (2.30)$$

Matriisien käsittelyssä hyödyllisiä Matlabin funktioita ovat mm. `eig`, `svd`, `pinv`, `orth`, `\` ja `null`. Huomautettakoon, että Matlabin SVD-algoritmi palauttaa kannat paitsi avaruudelle $\mathcal{R}(K)$ ja $\mathcal{N}(K)^\perp$, myös näiden ortogonaalisille komplementeille, joten palautetut matriisit ovat kuten kaavassa (2.28).

Sisältö:

1. Johdanto
2. Yleistetyt käänneiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 53 / 108

Takaisin

Full Screen

Sulje

Poistu



Esimerkki 2.13 Esitä K^\dagger aliavaruuksien $\mathcal{R}(K)$ ja $\mathcal{N}(K)^\perp$ ortonormaalien kantamatriisien U_1 ja V_1 avulla kuten kappaleessa 2.3.2. Ortonormaalisuuden perusteella saadaan helposti

$$\begin{aligned}g_1 &= U_1 U_1^T g, \\W &= K V_1 \quad \Gamma = W^T W \quad \Gamma c = \gamma = W^T g_1, \\c &= \Gamma^{-1} \gamma = \Gamma^{-1} W^T (U_1 U_1^T) g \quad \text{ja} \\f_1 &= V_1 c = \underbrace{V_1 (V_1^T K^T K V_1)^{-1} V_1^T K^T U_1 U_1^T}_{K^\dagger} g = K^\dagger g.\end{aligned}$$

□

Tehtäviä

Etusivu



Sivu 54 / 108

Takaisin

Full Screen

Sulje

Poistu

**Sisältö:**

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 55 / 108

Takaisin

Full Screen

Sulje

Poistu

3.1. Pseudoinverssin jatkuvuus

Tarkastellaan edelleen yhtälön

$$Kf = g \quad (3.1)$$

ratkeavuutta, kun $K : \mathcal{H}_1 \mapsto \mathcal{H}_2$ ja K on kompakti. Useissa tapauksissa yhtälöllä ei ole klassista ratkaisua, jolloin tarkastellaan miniminormista LS-ratkaisua. Yleistetty ratkaisu $f = K^\dagger g$ saattaa riippua g :stä epäjatkovasti, jolloin pienetkin datan virheet (kohina, pyöristysvirheet) voivat aiheuttaa suuria virheitä ratkaisuihin. Olkoon $\{\sigma_k, v_k, u_k\}$ operaattorin K singulaarisysteemi, joten ääretöndimensioisessa tapauksessa ($\dim \mathcal{N}(K)^\perp = \infty$)

$$K^\dagger g = \sum_{k=1}^{\infty} \frac{\langle g, u_k \rangle}{\sigma_k} v_k . \quad (3.2)$$

Kompakteille operaattoreille $\sigma_k \rightarrow 0$ ja piste 0 on **kasautumispiste**.

Jatkuvuus edellyttää lineaarikuvauksilta T , että $T(f_k - f) \rightarrow 0$, kun $f_k \rightarrow f$. Kun merkitään $w_k = f_k - f$, havaitaan, että $Tw_k \rightarrow 0$, kun $w_k \rightarrow 0$ ja $\|Tw_k\| \rightarrow 0$, kun $\|w_k\| \rightarrow 0$. Esimerkiksi



voidaan konstruoida jono $w_l = l^{-1}u_{k(l)}$, $l = 1, 2, \dots$. Kuvaus $k(l)$ valitsee kaikilla l indeksin k arvon siten, että $\sigma_k < l^{-1}$ ($\Rightarrow l^{-1}\sigma_k^{-1} > 1$). Tällöin $\|w_l\| = l^{-1}\|u_{k(l)}\| = l^{-1} \rightarrow 0$, kun $l \rightarrow \infty$. Mutta

$$\begin{aligned} \|K^\dagger w_l\| &= \|K^\dagger l^{-1}u_{k(l)}\| \\ &= l^{-1}\|K^\dagger u_{k(l)}\| \\ &= l^{-1}\left\|\sum_{n=1}^{\infty} \frac{\langle u_{k(l)}, u_n \rangle}{\sigma_n} v_n\right\| \\ &= l^{-1}\|\sigma_{k(l)}^{-1}v_{k(l)}\| \\ &= l^{-1}\sigma_{k(l)}^{-1}\|v_{k(l)}\| \\ &= l^{-1}\sigma_{k(l)}^{-1} \rightarrow a > 1, \quad \text{kun } l \rightarrow \infty. \end{aligned}$$

Siis $\|w_l\| \rightarrow 0$, mutta $\|K^\dagger w_l\| \not\rightarrow 0$, joten operaattorin K pseudoinverssi ei ole jatkuva.

Tihonov-regularisoinnin ideana on korvata tarkka huonosti asetettu ongelma hyvin asetetulla ongelmalla, jonka ratkaisu on jatkuva g :n suhteen. Tällä approksimatiivisen ongelman ratkaisulla approksimoidaan "tarkkaa" ratkaisua.

Yleistetty ratkaisu toteuttaa normaaliyhtälön

$$K^*Kf = K^*g. \quad (3.3)$$

Operaattori K^*K on itseadjungoitu ja kompakti, joten sen ominaisarvot ovat ei-negatiivisia. Siten operaattorin $K^*K + \alpha I$, $\alpha > 0$, ominaisarvoille λ_k pätee $\lambda_k \geq \alpha > 0$, joten operaattorilla $K^*K + \alpha I$ on rajoitettu käänteiskuvaus (laskuharjoitukset), jolloin ongelman

$$(K^*K + \alpha I)f_\alpha = K^*g \quad (3.4)$$

ratkaisu on stabiili (hyvin asetettu ongelma). Tämän ratkaisua

$$f_\alpha = (K^*K + \alpha I)^{-1}K^*g \quad (3.5)$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 56 / 108

Takaisin

Full Screen

Sulje

Poistu



kutsutaan ratkaisun $f = K^\dagger g$ *Tihonov-approksimaatioksi* ja parametria α *regularisointiparametriksi*. Ratkaisu voidaan yksinkertaistaa matriisimuotoon, jolloin

$$f_\alpha = (K^T K + \alpha I)^{-1} K^T g.$$

3.2. SVD-muoto

Tihonov-approksimaatio voidaan esittää singulaariarvohajotelman avulla seuraavasti. Olkoon $\{\sigma_k; v_k, u_k\}$ operaattorin K singulaarisysteemi. Nyt $\alpha f_\alpha = K^* g - K^* K f_\alpha$, joten $f_\alpha \in \mathcal{R}(K^*) = \mathcal{N}(K)^\perp$. Siten f_α voidaan esittää muodossa

$$f_\alpha = \sum_{k=1}^{\infty} \langle f_\alpha, v_k \rangle v_k .$$

Vastaavasti

$$\begin{aligned} K^* g &= \sum_{k=1}^{\infty} \langle K^* g, v_k \rangle v_k \\ &= \sum_{k=1}^{\infty} \langle g, K v_k \rangle v_k \\ &= \sum_{k=1}^{\infty} \sigma_k \langle g, u_k \rangle v_k . \end{aligned} \tag{3.6}$$

Edelleen muodosta (3.4) saadaan

$$\begin{aligned} K^* g &= (K^* K + \alpha I) f_\alpha \\ &= (K^* K + \alpha I) \sum_{k=1}^{\infty} \langle f_\alpha, v_k \rangle v_k \end{aligned} \tag{3.7}$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 57 / 108

Takaisin

Full Screen

Sulje

Poistu



$$\begin{aligned}
&= \sum_{k=1}^{\infty} \langle f_{\alpha}, v_k \rangle (K^* K + \alpha I) v_k \\
&= \sum_{k=1}^{\infty} \langle f_{\alpha}, v_k \rangle \underbrace{(K^* K v_k + \alpha v_k)}_{\sigma_k^2 v_k} .
\end{aligned} \tag{3.8}$$

Tällöin esitysten (3.6) ja (3.8) perusteella saadaan

$$\begin{aligned}
\sum_{k=1}^{\infty} (\sigma_k^2 + \alpha) \langle f_{\alpha}, v_k \rangle v_k &= \sum_{k=1}^{\infty} \sigma_k \langle g, u_k \rangle v_k \Leftrightarrow \\
\langle f_{\alpha}, v_k \rangle &= \frac{\sigma_k}{\sigma_k^2 + \alpha} \langle g, u_k \rangle \Leftrightarrow \\
f_{\alpha} &= \sum_{k=1}^{\infty} \frac{\sigma_k}{\sigma_k^2 + \alpha} \langle g, u_k \rangle v_k .
\end{aligned} \tag{3.9}$$

Esimerkki 3.1 Äärellisdimensioisessa tapauksessa olkoon $K = USV^T$. Seuraava Matlab-fragmentti laskee Tihonov-regularisoidun ratkaisun

```
[U,S,V]=svd(K); % lasketaan matriisin K singulaarisysteemi
sd=diag(S);      % matriisin S diagonaalialkiot vektoriksi
sdli=diag(sd./(sd.^2+alpha)); % lasketaan kertoimet ja
fl = V*sdli*U'*g;
```

Määritellään “suodinkertoimet” (filter factors) $q(\sigma_k)$ ja kirjoitetaan

$$f_{\text{reg}} = \sum_{k=1}^{\infty} q(\sigma_k) \langle g, u_k \rangle v_k .$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 58 / 108

Takaisin

Full Screen

Sulje

Poistu



Useat regularisointimenetelmät voidaan esittää määrittelemällä suodinkertoimet sopivasti. Esimerkiksi

$$\begin{aligned} \text{pseudoinverssi} & q(\sigma_k) = \sigma_k^{-1}, \\ \text{katkaistu pseudoinverssi} & q(\sigma_k) = \begin{cases} \sigma_k^{-1}, & 1 \leq k \leq r \\ 0, & k > r \end{cases} \quad \text{ja} \\ \text{Tihonov-regularisointi} & q(\sigma_k) = \frac{\sigma_k}{\sigma_k^2 + \alpha}. \end{aligned}$$

3.3. Konvergenssi ja kohinainen data

Tihonov-approksimaatiota voi edellä esitetyn perusteella pitää *ad hoc* -menetelmänä. Tarkastellaan seuraavassa approksimaation konvergenssia, eli pätee $f_\alpha \rightarrow f$, kun $\alpha \rightarrow 0$? Mikäli näin on eli esimerkiksi, jos K^\dagger on stabiili, voidaan periaatteessa muodostaa jono ratkaisuja $f_{\alpha_1}, f_{\alpha_2}, \dots$, joka suppenee kohti ratkaisua f , kun $\alpha_k \rightarrow 0$. Käänteisesti ajateltuna, jos $Kf = g$ ja $g \in \mathcal{R}(K)$ on virheetön, niin

$$\sigma_k^{-1} \langle g, u_k \rangle \rightarrow 0, \quad \text{kun } k \rightarrow \infty.$$

Miniminnormiratkaisulle saatiin aikaisemmin

$$f = K^\dagger g = \sum_{k=1}^{\infty} \sigma_k^{-1} \langle g, u_k \rangle v_k. \quad (3.10)$$

Tarkastellaan näiden ratkaisujen erotusta

$$\begin{aligned} f_\alpha - K^\dagger g &= \sum_{k=1}^{\infty} \frac{\sigma_k}{\sigma_k^2 + \alpha} \langle g, u_k \rangle v_k - \sum_{k=1}^{\infty} \frac{1}{\sigma_k} \langle g, u_k \rangle v_k \\ &= \sum_{k=1}^{\infty} \left(\frac{\sigma_k}{\sigma_k^2 + \alpha} - \frac{1}{\sigma_k} \right) \langle g, u_k \rangle v_k \\ &= \sum_{k=1}^{\infty} \frac{-\alpha}{\sigma_k(\sigma_k^2 + \alpha)} \langle g, u_k \rangle v_k. \end{aligned}$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 59 / 108

Takaisin

Full Screen

Sulje

Poistu



Erotuksen normin neliö

$$\|f_\alpha - K^\dagger g\|^2 = \sum_{k=1}^{\infty} \left(\frac{-\alpha}{\sigma_k(\sigma_k^2 + \alpha)} \right)^2 |\langle g, u_k \rangle|^2 .$$

Jotta voidaan tarkastella oikean puolen konvergenssia, kun $\alpha \rightarrow 0$ ja vaihtaa raja-arvon ja summauksen järjestyks, on todettava, että oikea puoli on rajoitettu. Tämä nähdään seuraavasti. Koska $\sigma_k > 0$ kaikille k ja

$$\frac{\alpha}{\sigma_k(\sigma_k^2 + \alpha)} = \frac{1}{\underbrace{\sigma_k(\sigma_k^2/\alpha + 1)}_{\geq 1}} \leq \frac{1}{\sigma_k} ,$$

saadaan

$$\left(\frac{\alpha}{\sigma_k(\sigma_k^2 + \alpha)} \right)^2 |\langle g, u_k \rangle|^2 \leq \sigma_k^{-2} |\langle g, u_k \rangle|^2$$

kaikille k . Koska g on kiinteä, voi $K^\dagger g$ olla hyvinkin suuri, mutta $\|K^\dagger g\| < \infty$. Koska

$$\infty > \|K^\dagger g\|^2 = \sum_{k=1}^{\infty} \sigma_k^{-2} |\langle g, u_k \rangle|^2 ,$$

voidaan edellä mainittujen operaatioiden järjestyks vaihtaa, joten

$$\begin{aligned} \lim_{\alpha \rightarrow 0} \|f_\alpha - K^\dagger g\|^2 &= \lim_{\alpha \rightarrow 0} \sum_{k=1}^{\infty} \left(\frac{-\alpha}{\sigma_k(\sigma_k^2 + \alpha)} \right)^2 |\langle g, u_k \rangle|^2 \\ &= \sum_{k=1}^{\infty} |\langle g, u_k \rangle|^2 \underbrace{\lim_{\alpha \rightarrow 0} \left(\frac{-\alpha}{\sigma_k(\sigma_k^2 + \alpha)} \right)^2}_{=0 \text{ kaikille } k} \\ &= 0 . \end{aligned}$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 60 / 108

Takaisin

Full Screen

Sulje

Poistu



Koska operaattori $(K^*K + \alpha I)^{-1}K^*$ on rajoitettu, Tihonov-approksimaatio f_α riippuu jatkuvasti datasta g kaikille kiinnitetyille $\alpha > 0$.

Tarkastellaan seuraavaksi Tihonov-approksimaatiota tapauksessa, jossa data on kohinaainen. Olkoon havaitun datan g^δ virheen normi korkeintaan δ , eli

$$\|g - g^\delta\| \leq \delta,$$

missä (toistaiseksi) oletetaan, että δ tunnetaan. Nyt Tihonov-approksimaatio vastaa kohinaista dataa, eli

$$f_\alpha^\delta = (K^*K + \alpha I)^{-1}K^*g^\delta.$$

Tarkastellaan erotusta

$$f_\alpha^\delta - f_\alpha = (K^*K + \alpha I)^{-1}K^*(g^\delta - g).$$

Havaitaan, että $(K^*K + \alpha I)^{-1}K^* = K^*(KK^* + \alpha I)^{-1}$ (matriisin kääntölemma), jolloin saadaan

$$\begin{aligned} \|f_\alpha^\delta - f_\alpha\|^2 &= \langle K^*(KK^* + \alpha I)^{-1}(g^\delta - g), K^*(KK^* + \alpha I)^{-1}(g^\delta - g) \rangle \\ &= \langle (KK^* + \alpha I)^{-1}(g^\delta - g), KK^*(KK^* + \alpha I)^{-1}(g^\delta - g) \rangle \\ &\leq \underbrace{\|KK^*(KK^* + \alpha I)^{-1}\|}_{\leq 1} \langle (KK^* + \alpha I)^{-1}(g^\delta - g), g^\delta - g \rangle. \end{aligned}$$

Toisaalta $\|KK^* + \alpha I\| \geq \alpha$ ja $\|(KK^* + \alpha I)^{-1}\| \leq \alpha^{-1}$, joten (harjoitustehtävä)

$$\|f_\alpha^\delta - f_\alpha\|^2 \leq \alpha^{-1} \langle g^\delta - g, g^\delta - g \rangle = \alpha^{-1} \underbrace{\|g^\delta - g\|^2}_{\delta^2} \Leftrightarrow$$

$$\|f_\alpha^\delta - f_\alpha\| \leq \delta / \sqrt{\alpha}.$$

Siten kaikille kiinteille $\delta > 0$ ratkaisun virhe voi kasvaa rajatta, kun regularisointiparametri $\alpha \rightarrow 0$.

Perinteisesti keskeinen ongelma regularisoinnissa onkin, kuinka regularisointiparametri tulisi valita (vastaten tunnettua δ) siten, että kohinaista dataa vastaava ratkaisu suppenee kohti

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 61 / 108

Takaisin

Full Screen

Sulje

Poistu



tarkkaa ratkaisua, kun $\delta \rightarrow 0$. Toisin sanoen, vaatimuksena on, että

$$\alpha(\delta) \rightarrow 0 \quad \text{ja} \quad f_{\alpha(\delta)}^{\delta} \rightarrow K^{\dagger}g, \quad \text{kun} \quad \delta \rightarrow 0 .$$

Tällöin regularisointimenetelmää kutsutaan *säännölliseksi* (*regular algorithm*). Ongelmaksi jää yleensä se, että virheen normia δ ei tiedetä tarkasti ja että virhe ei useinkaan ole jakautunut tasaisesti havaintoihin, jolloin tämän tyyppiset tarkastelut eivät välttämättä ole relevantteja.

Missä sovelluksissa voidaan ajatella, että δ on hyvin pieni? Entä miten virheen $v = g - g^{\delta}$ “rakenne” tulisi ottaa huomioon, eli onko virhe esimerkiksi “tasaisesti” jakautunut dataan g ? Jos näin ei ole, onko ylipäättään oleellista ainoastaan tarkastella regularisointiparametrimin α suuruutta? Lisäksi kannattaa pohtia, kuinka hyvin δ tunnetaan. Kuinka “herkkä” ratkaisu on, kun $\delta^ = \delta + \tilde{\delta}$?*

3.4. Variaatiomuoto

Tarkastellaan neliöllistä funktionaalia

$$F_{\alpha}(f) = \|Kf - g^{\delta}\|^2 + \alpha\|f\|^2 . \quad (3.11)$$

Jatkoa silmällä pitäen tulee funktionaalin $F_{\alpha}(f)$ toinen osa lukea kuten $\alpha\|f\|^2$. Tätä termiä kutsutaan tavallisesti *sivurajoitteeksi*, vaikka kysymyksessä ei ole tavanomainen rajoite. Funktionaalin ensimmäinen osa edustaa etäisyyttä miniminormiratkaisuun eli residuaalia. Sen sijaan, että miniminormiratkaisu $f = K^{\dagger}g^{\delta}$ olisi yksikäsitteinen, on ratkaisu käytännössä mikä tahansa vektori miniminormiratkaisun ja operaattorin K oleellisen nolla-avaruuden määräämässä monistossa. Termillä oleellinen viitataan tässä valintaan, joka koskee operaattorin K singulaariarvojen tulkitsemista käytännössä nolliksi. Minimoitavan funktionaalin (3.11) toinen osa taas “ajaa” vastaavaa LS-ratkaisua normiltaan pienempään suuntaan, eli (todella) löysästi sanoen pyrkii valitsemaan edellä mainitun moniston alkioista jonkin normiltaan pienistä.

Tarkastellaan LS-ongelman

$$f_{\alpha} = \text{sol} \min_f F_{\alpha}(f) , \quad F_{\alpha} : \mathcal{H}_1 \mapsto \mathbb{R}$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 62 / 108

Takaisin

Full Screen

Sulje

Poistu



ratkaisua $f \in \mathcal{H}_1$, kun $K : \mathcal{H}_1 \mapsto \mathcal{H}_2$. Määritellään normi karteesisessa tuloavaruudessa $\mathcal{H}_1 \times \mathcal{H}_2$, kun $u \in \mathcal{H}_1$ ja $v \in \mathcal{H}_2$ siten, että

$$\|(u, v)\|_{\mathcal{H}_1 \times \mathcal{H}_2}^2 = \|v\|_{\mathcal{H}_2}^2 + \alpha \|u\|_{\mathcal{H}_1}^2 .$$

Nyt operaattorin K graafi (f, Kf) , $f \in \mathcal{H}_1$, on suljettu ja konvekksi, joten on olemassa $f \in \mathcal{H}_1$ siten, että $\|(f, Kf)\|$ saa minimin. Edelleen (muuttujanvaihdon avulla) nähdään, että siten myös funktionaalilla (3.11) on minimi, joka on pisteen $(0, g^\delta)$ pienin neliöetäisyys operaattorin K graafista.

Siis funktionaalin (3.11) minimoivalle alkioille z pätee

$$\frac{d}{dt} \{ \|K(z + tw) - g^\delta\|^2 + \alpha \|z + tw\|^2 \} \Big|_{t=0} = 0$$

kaikille $w \in \mathcal{H}_1$ ja $t \in \mathbb{R}$. Siten

$$\begin{aligned} & \frac{d}{dt} \{ \|K(z + tw) - g^\delta\|^2 + \alpha \|z + tw\|^2 \} \Big|_{t=0} \\ &= \frac{d}{dt} \{ \langle K(z + tw) - g^\delta, K(z + tw) - g^\delta \rangle + \alpha \langle z + tw, z + tw \rangle \} \Big|_{t=0} . \end{aligned}$$

Kun $t \in \mathbb{R}$, niin sisätulon bilineaarisuuden perusteella

$$\begin{aligned} \frac{d}{dt} \langle z + tu, v \rangle \Big|_{t=0} &= \frac{d}{dt} \{ \langle z, v \rangle + t \langle u, v \rangle \} \Big|_{t=0} \\ &= \langle u, v \rangle \quad \text{ja} \\ \frac{d}{dt} \langle z + tu, z + tu \rangle \Big|_{t=0} &= \frac{d}{dt} \{ \langle z, z \rangle + t \langle z, u \rangle + t \langle u, z \rangle + t^2 \langle u, u \rangle \} \Big|_{t=0} \\ &= 2 \langle z, u \rangle , \end{aligned}$$

joten

$$\begin{aligned} \langle Kz - g^\delta, Kw \rangle + \alpha \langle z, w \rangle &= 0 \quad \text{ja} \\ \langle (K^*K + \alpha I)z - K^*g^\delta, w \rangle &= 0 \end{aligned} \tag{3.12}$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 63 / 108

Takaisin

Full Screen

Sulje

Poistu



kaikille $w \in \mathcal{H}_1$. Näin ollen on oltava

$$(K^*K + \alpha I)z = K^*g^\delta .$$

Siten funktionaalin (3.11) minimille pätee

$$f_\alpha^\delta = (K^*K + \alpha I)^{-1}K^*g^\delta ,$$

eli se on sama kuin miniminormiratkaisun f Tihonov-approksimaatio. Käytännössä Tihonov-approksimaatiota ei edellisestä muodosta juuri koskaan lasketa. Käytännön menetelmä esitetään seuraavassa kappaleessa.

3.5. Yleistetty Tihonov-regularisointi

Tarkastellaan funktionaalin (3.11) yleistystä

$$F_\alpha(f) = \|Kf - g^\delta\|^2 + \alpha\|Lf\|^2 .$$

Tätä vastaava variaatiomuoto (harjoitustehtävä) on

$$\langle (K^*K + \alpha L^*L)z - K^*g^\delta, w \rangle = 0 .$$

Sivurajoite Lf voidaan vielä korvata muodolla $L(f - f_*)$, jonka merkityksestä enemmän myöhemmin. Näin saadaan *yleistetty Tihonov-ongelma*

$$f_\alpha^\delta = \operatorname{sol} \min_f \{ \|Kf - g^\delta\|^2 + \alpha\|L(f - f_*)\|^2 \} . \quad (3.13)$$

Mikäli tarpeellista, voidaan suorittaa muuttujanvaihto $u = f - f_*$, jolloin ongelma saadaan muotoon

$$u = \operatorname{sol} \min_u \{ \|Ku - g'\|^2 + \alpha\|Lu\|^2 \} , \quad (3.14)$$

missä $g' = g^\delta - Kf_*$. Edelleen, jos L on säännöllinen, merkitään $Lu = v$ ja $u = L^{-1}v$, jolloin saadaan

$$v = \operatorname{sol} \min_v \{ \|KL^{-1}v - g'\|^2 + \alpha\|v\|^2 \} .$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 64 / 108

Takaisin

Full Screen

Sulje

Poistu



Tarkastellaan muotoa (3.13) äärellisdimensioisessa tapauksessa. Olkoon nyt $K \in \mathbb{R}^{N_1 \times M}$ ja $L \in \mathbb{R}^{N_2 \times M}$. Merkitään $e_1 = Kf - g^\delta \in \mathbb{R}^{N_1}$, $e_2 = \sqrt{\alpha}L(f - f_*) \in \mathbb{R}^{N_2}$ ja $e = [e_1^T, e_2^T]^T$. Ongelma (3.14) voidaan nyt kirjoittaa muotoon

$$f_\alpha^\delta = \text{sol min}_f \|\tilde{K}f - \tilde{g}\|^2 = \text{sol min}_f \|e\|^2 \quad (3.15)$$

$$= \text{sol min}_f \{\|e_1\|^2 + \|e_2\|^2\} , \quad (3.16)$$

missä

$$\tilde{K} = \begin{pmatrix} K \\ \sqrt{\alpha}L \end{pmatrix} \quad \text{ja} \quad \tilde{g} = \begin{pmatrix} g^\delta \\ \sqrt{\alpha}Lf_* \end{pmatrix} . \quad (3.17)$$

Siten Tihonov-regularisoidun ratkaisun määrääminen palautuu muunnetun (lisätyn) LS-ongelman ratkaisun määräämiseen! Jos ongelman dimensio ei ole suuri ja matriisin \tilde{K} sarakeranki on täysi¹, voidaan periaatteessa kirjoittaa

$$f_\alpha^\delta = (\tilde{K}^* \tilde{K})^{-1} \tilde{K}^* \tilde{g} . \quad (3.18)$$

Käytännössä tällöinkin ratkaisu lasketaan esimerkiksi QR-hajotelman avulla, eikä käänteismatriisia $(\tilde{K}^* \tilde{K})^{-1}$ jouduta laskemaan.

Useissa tapauksissa, kuten esimerkiksi osittaisdifferentiaaliyhtälöihin liittyvissä parametriestimointi- ja säätöongelmissa, \tilde{K} on hyvin suuri ja lisäksi harva matriisi. Tällöin ratkaisu voidaan määrätä iteratiivisesti esimerkiksi jollakin kiintopisteiteraatiolla, katso Luku 4. Hyvin pienidimensioisissa ongelmissa voidaan käyttää muotoa 3.18. Seuraava Matlab-fragmentti laskee ratkaisun QR-hajotelman avulla.

```
fda=[K;sqrt(alpha)*L]\[gd;sqrt(alpha)*L*fst];
```

HUOMAUTUS. Kuinka voidaan tarkastaa, onko $\mathcal{N}(K) \cap \mathcal{N}(L) = \{0\}$? Määrätään operaattoreiden K ja L nolla-avaruuksille kannat (esimerkiksi SVD:n avulla) $\mathcal{N}(K) = \text{span}\{w_k\}$ ja $\mathcal{N}(L) = \text{span}\{v_k\}$. Projisoidaan vektorit v_k yksitellen nolla-avaruudelle $\mathcal{N}(K)$ ja määritellään

$$q_\ell = \{v_\ell \mid v_\ell = P_{\mathcal{N}(K)} v_\ell\} .$$

¹Jos näin ei ole, on voimassa $\mathcal{N}(K) \cap \mathcal{N}(L) \neq \{0\}$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 65 / 108

Takaisin

Full Screen

Sulje

Poistu



Nyt selvästi jokainen $q_\ell \in \mathcal{N}(K) \cap \mathcal{N}(L)$. \square

3.5.1. Yleistetty singulaariarvohajotelma

Olkoon $K \in \mathbb{R}^{N_1 \times M}$ ja $L \in \mathbb{R}^{N_2 \times M}$, missä yksinkertaisuuden vuoksi (tässä) oletetaan, että $N_1 \geq M \geq N_2$. Matriisiparin (K, L) *yleistetty singulaariarvohajotelma* (*generalized singular value decomposition, GSVD*) on

$$K = U \begin{pmatrix} \Psi & 0 \\ 0 & I_{M-N_2} \end{pmatrix} X^{-1}, \quad L = V(\Xi 0)X^{-1}, \quad (3.19)$$

missä $U = [u_1, \dots, u_{N_1}] \in \mathbb{R}^{N_1 \times N_1}$ ja $V = [v_1, \dots, v_{N_2}] \in \mathbb{R}^{N_2 \times N_2}$ ovat ortonormaaleja, $X = [x_1, \dots, x_M] \in \mathbb{R}^{M \times M}$ on säännöllinen, $\Psi = \text{diag}(\psi_1, \dots, \psi_{N_2})$ ja $\Xi = \text{diag}(\xi_1, \dots, \xi_{N_2})$. Lisäksi $0 \leq \psi_1 \leq \dots \leq \psi_{N_2} \leq 1$, $1 \geq \xi_1 \geq \dots \geq \xi_{N_2} \geq 0$ ja $\psi_k^2 + \xi_k^2 = 1$, kun $k = 1, \dots, N_2$.

Matriisiparin (K, L) *yleistetyt singulaariarvot* ovat luvut

$$\gamma_k = \psi_k / \xi_k, \quad k = 1, \dots, N_2$$

ja $0 \leq \gamma_1 \leq \gamma_2 \leq \dots \leq \gamma_{N_2}$. Jos V ei ole täysiasteinen, $\xi_k = 0$ kun $k > \text{rank}(V)$ ja yleistetty singulaariarvo $\gamma_k = \infty$.

Regularisoitu ratkaisu voidaan kirjoittaa GSVD:n avulla muotoon

$$f_\alpha(L) = \sum_{k=1}^p \frac{\gamma_k^2}{\psi_k(\gamma_k^2 + \alpha)} \langle g, u_k \rangle x_k + \sum_{k=p+1}^M \langle g, u_k \rangle x_k, \quad (3.20)$$

missä $p = \text{rank}(L)$, eli $p = N_2$, jos $\mathcal{N}(L) = \{0\}$. Yhtälön (3.20) oikean puolen toinen summa ei riipu regularisointiparametrin α arvosta.

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 66 / 108

Takaisin

Full Screen

Sulje

Poistu



3.6. Sivurajoituksen merkitys

3.6.1. Deterministinen tulkinta

Sivurajoitteessa $\|L(f - f_*)\|$ ei ole kysymys varsinaisesta rajoitteesta, mutta jos $\alpha \rightarrow \infty$, saadaan regularisoiduksi ratkaisuksi yhtälörajoitetun LS-ongelman

$$f = \text{sol} \quad \inf_{L(f - f_*)=0} \|Kf - g\|^2 \quad (3.21)$$

ratkaisu, missä $L(f - f_*) = 0$ on yhtälön rajoite.

Esimerkki 3.2 Oletetaan esimerkiksi, että $L(f - f_*) = 0$, $V = [v_1, \dots, v_r]_{N \times r}$ ja $\mathcal{N}(L) = \text{span}\{v_1, \dots, v_r\}$. Tällöin

$$f - f_* = Vc, \quad c \in \mathbb{R}^r$$

ja

$$\begin{aligned} Kf - g &= K(Vc + f_*) - g \\ &= \tilde{K}c - \tilde{g}. \end{aligned}$$

Nyt voidaan ratkaista c ja sen jälkeen laskea $f = Vc + f_*$.

HUOMAUTUS. Miten käy, jos $L \in \mathbb{R}^{M \times N}$ ja $\mathcal{N}(L) = \{0\}$? Kun

$$\begin{aligned} M > N &: f = f_*, \\ M = N &: f = f_* \quad \text{ja} \\ M < N &: \text{ei mahdollinen.} \end{aligned}$$

Useimmissa tapauksissa sopivalla regularisointioperaattorilla on ei-triviaali nolla-avaruus $\mathcal{N}(L)$.
□

Tarkastellaan sivurajoituksen merkitystä ongelman (3.21) ratkaisulle ja asetetaan $f_* = 0$. Tapauksessa $L = I$ regularisointiparametrin kasvattaminen eli $\alpha \rightarrow \infty$ aiheuttaa, että ratkaisu

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 67 / 108

Takaisin

Full Screen

Sulje

Poistu



$f_\alpha(I) \rightarrow 0$. Yleisemmässä tapauksessa operaattorille L pätee $f_\alpha(L) \rightarrow f_1 \in \mathcal{N}(L)$. Tämä on osoitettavissa GSVD-esityksen (3.20) avulla.

Tarkastellaan spatiaalista yksiulotteista ongelmaa, jossa $t = (t_1, \dots, t_N)$, $t_{k+1} - t_k = h$, $f = f(t)$ ja $f_1 = f(t_1), \dots, f_N = f(t_N)$. Oletetaan, että L vastaa diskreetoitua ensimmäisen kertaluvun derivaattaoperaattoria, eli

$$L = D_1 = \begin{pmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \end{pmatrix} \in \mathbb{R}^{(N-1) \times N}. \quad (3.22)$$

Tämän operaattorin nolla-avaruus käsittää kaikki vakiovektorit, eli $\mathcal{N}(L) = \{c(1, \dots, 1) \in \mathbb{R}^N, c \in \mathbb{R}\}$, joten ratkaisun $f_\alpha(D_1)$ raja-arvo on vakiovektori (havainnot vastaavat funktiosta $f(t) = a$ saatuja arvoja).

Olkoon D_m m :n kertaluvun differenssiä vastaava matriisi, eli esimerkiksi

$$D_2 = \begin{pmatrix} -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & \ddots & \ddots & \ddots \\ & & & -1 & 2 & -1 \end{pmatrix} \in \mathbb{R}^{(N-2) \times N} \quad \text{ja}$$

$$D_3 = \begin{pmatrix} -1 & 3 & -3 & 1 & & \\ & -1 & 3 & -3 & 1 & \\ & & \ddots & \ddots & \ddots & \ddots \\ & & & -1 & 3 & -3 & 1 \end{pmatrix} \in \mathbb{R}^{(N-3) \times N}.$$

Edellisen perusteella raja-arvoille saadaan siten

$$f_\alpha(D_2) = \text{sol} \min_f \|Kf - g\|^2 + \alpha \|D_2 f\|^2 \rightarrow f_2, \quad \text{kun } \alpha \rightarrow \infty \quad \text{ja}$$

$$f_\alpha(D_3) = \text{sol} \min_f \|Kf - g\|^2 + \alpha \|D_3 f\|^2 \rightarrow f_3, \quad \text{kun } \alpha \rightarrow \infty,$$

Sisältö:

1. Johdanto
2. Yleistetyt käännekuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 68 / 108

Takaisin

Full Screen

Sulje

Poistu



missä $f_2 \in \mathcal{N}(D_2)$ (havainnot vastaavat funktiosta $f(t) = a + bt$ saatuja arvoja) ja $f_3 \in \mathcal{N}(D_3)$ (havainnot vastaavat funktiosta $f(t) = a + bt + ct^2$ saatuja arvoja).

Näitä sivurajoitteiden raja-arvojen nolla-avaruuksien karakterisoimia ratkaisuja kutsutaan joskus *asymptoottiseksi prioreiksi*. Siten tapauksessa $L = I$ asymptoottinen priorin on $f \equiv 0$. Seuraavan **MATLAB-esimerkin** avulla voit tarkastella regularisointioperaattorin ja -parametrin vaikutusta estimaattiin.

Eräs tapa “koodata” a priori-oletukset (etukäteen tiedossa oleva tieto ratkaisusta) sivurajoitteeseen on seuraava. Olkoon olemassa hyvä arvaus monistosta $\mathcal{V}_* = \{f = f_* + v \mid v \in \mathcal{V}\}$, jossa ratkaisun f_α^δ oletetaan olevan. Olkoon $\mathcal{V} = \text{span}\{v_1, \dots, v_r\}$ ja $V = (v_1, \dots, v_r)$, missä vektorijoukko $\{v_k\}$ on ortonormaali. Muodostetaan sivurajoite

$$L(f - f_*) = (I - VV^T)(f - f_*) = (I - P_V)(f - f_*) .$$

Nyt jos $f \in \mathcal{V}_*$, niin $(I - P_V^T)(f - f_*) = 0$, joten kyseinen sivurajoite “vetää” regularisointua ratkaisua f_α^δ kohti monistoa \mathcal{V}_* . Regularisoidun ratkaisun etäisyydelle monistosta \mathcal{V}_* pätee $\text{dist}(f_\alpha^\delta, \mathcal{V}_*) \rightarrow 0$, kun $\alpha \rightarrow \infty$ (**harjoitustehtävä**). Useissa tapauksissa moniston \mathcal{V}_* konstruoiminen saattaa osoittautua hyvin työlääksi.

Esimerkki 3.3 *Olkoon \mathcal{V}_* toisen asteen polynomien joukko, eli $\mathcal{V}_* = \{f(t) \in \mathcal{F}([0, 1], \mathbb{R}) \mid f(t) = c_0 + c_1t + c_2t^2, f(0) = 1\}$. Olkoon $f \in \mathcal{V}_*$, eli $f(t) = c_0 + c_1t + c_2t^2$. Koska $f(0) = 1 \Rightarrow c_0 = 1$. Olkoon $f_*(t) \equiv 1$, jolloin $f - f_* \in \mathcal{V} = \text{span}\{t, t^2\}$. Tehdään \mathcal{V} :lle ortonormaalikanta Gram-Schmidtin -ortogonalisointimenetelmällä:*

Olkoon $u_1 = t$ ja $u_2 = t^2$. Nyt

$$\begin{aligned} \tilde{v}_1 &= u_1 = t , \\ v_1 &= \|\tilde{v}_1\|^{-1} \tilde{v}_1 = \left(\int_0^1 t^2 dt \right)^{-1/2} t = \frac{1}{\sqrt{3}} t , \\ \tilde{v}_2 &= u_2 - \langle u_2, v_1 \rangle v_1 \\ &= t^2 - \left(\int_0^1 t^2 \frac{1}{\sqrt{3}} t dt \right) \frac{1}{\sqrt{3}} t = t^2 - \frac{1}{12} t \quad \text{ja} \end{aligned}$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 69 / 108

Takaisin

Full Screen

Sulje

Poistu



$$\begin{aligned}v_2 &= \|\tilde{v}_2\|^{-1}\tilde{v}_2 = \left(\int_0^1 t^2 - \frac{1}{12}tdt\right)^{-1/2}\left(t^2 - \frac{1}{12}t\right) \\ &= \sqrt{\frac{24}{7}}\left(t^2 - \frac{1}{12}t\right).\end{aligned}$$

Nyt f :n projektio \mathcal{V} :lle voidaan kirjoittaa muodossa

$$P_{\mathcal{V}}f = \sum_{k=1}^2 \langle f, v_k \rangle v_k.$$

Edelleen funktio $y(t) = P_{\mathcal{V}}(f - f_*) = \sum_{k=1}^2 \langle f - f_*, v_k \rangle v_k = \dots$ on toisenasteen polynomi, jolle $y(0) = 0$. Jos $f(x)$ on paraabeli, jolle $f(0)=0$, niin $y(x)=f(x)$.

Joissakin tapauksissa on kuitenkin mahdollista konstruoida odotettavissa olevia ratkaisuja kuvaava opetusjoukko $\{f_k, k = 1, \dots, r_1\}$. Tällöin voidaan edetä seuraavasti. Muodostetaan suureet

$$\begin{aligned}f_* &= r_1^{-1} \sum_{k=1}^{r_1} f_k, \\ q_k &= f_k - f_* \quad \text{jä} \\ \Gamma_* &= r_1^{-1} \sum_{k=1}^{r_1} q_k q_k^T.\end{aligned}$$

Tavoitteena on nyt määrätä (pienidimensioinen) aliavaruus \mathcal{V} siten, että vektoreiden q_k etäisyys tästä aliavaruudesta on pieni. Mikäli pienuutta on tarkoituksenmukaista mitata neliöetäisyydellä, \mathcal{V} määrätään matriisiin Γ_* ominaisarvohajoittelman $\Gamma_* = W\Lambda W^T$ avulla, missä

$$\Lambda = \text{diag}(\lambda_1, \dots, \lambda_N) \quad \text{jä} \quad W = [w_1, \dots, w_N].$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtävät

Etusivu



Sivu 70 / 108

Takaisin

Full Screen

Sulje

Poistu



Asetetaan $\mathcal{V} = \text{span}\{w_k, k = 1, \dots, r\}$ ja $V = (w_1, \dots, w_r)$, kun $r < r_1, N$. Kuten edelläkin, menetelmän toimivuus perustuu havaintoon

$$P_{\mathcal{V}^\perp} = I - VV^T,$$

joten $(I - VV^T)(f - f_*)$ on sitä pienempi, mitä pienempi vektorin f etäisyys monistosta \mathcal{V}_* on.

3.7. Epälineaariset käänteisongelmat

Tässä kappaleessa tarkastellaan epälineaaristen käänteisongelmien ratkaisua Tihonov-regularisoinnin avulla. Ongelma on muotoa

$$f = \text{sol} \min_f \{ \|g - F(f)\|^2 + \alpha \|L(f - f_*)\|^2 \}, \quad (3.23)$$

missä kuvaus F on epälineaarinen. Tarkastellaan yksinkertaisuuden vuoksi äärellisdimensioista ongelmaa $F : \mathbb{R}^N \mapsto \mathbb{R}^M$. Oletetaan, että kuvaus F on **Fréchet -derivoituva** kaikkialla, jolloin voidaan kirjoittaa

$$\begin{aligned} F(f) &= F(f_0) + F'(f_0)(f - f_0) + \mathcal{O}(\|f - f_0\|^2) \quad \text{ja} \\ F(f_0 + \delta f) &= F(f_0) + F'(f_0)\delta f + \mathcal{O}(\|\delta f\|^2), \end{aligned}$$

missä $\delta f = f - f_0$. Jos F on lineaarinen kuvaus, se on muotoa $F(f) = Kf$, jolloin $F'(f) \equiv K$ kaikille f .

Newtonin menetelmän ideana on ratkaista ongelma (3.23) iteratiivisesti siten, että kyseistä funktionaaliala approksimoidaan jossakin pisteessä f :n suhteen neliöllisellä funktionaalilla, jota vastaava ratkaisu on \bar{f} . Tämän jälkeen funktionaalin neliöllinen approksimaatio suoritetaan pisteessä \bar{f} ja toistetaan menettely. Jos tämä linearisointujen ongelmien ratkaisujen jono suppenee, se suppenee alkuperäisen ongelman paikalliseen ääriarvokohtaan. Gauss-Newton-menetelmä on edellisen modifikaatio, jossa Newtonin menetelmässä esiintyvää Hessin matriisia (kyseisen funktionaalin toisten derivaattojen muodostama matriisi) approksimoidaan matriisilla $(F')^T F'$. Gauss-Newton-menetelmään päästään myös suoraan määräämällä linearisoitua kuvausta vastaava LS-ratkaisu.

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 71 / 108

Takaisin

Full Screen

Sulje

Poistu



Pyritään rekursiiviseen menetelmään, jossa iteraatioiden jono $f^{(j)}$ on muotoa

$$f^{(j+1)} = f^{(j)} + \delta f^{(j)} ,$$

missä tehtävänä on nyt määrätä *päivitykset* $\delta f^{(j)}$ siten, että $f^{(j+1)}$ on pisteessä $f^{(j)}$ linearisoidun ongelman ratkaisu. Linearisoitua kuvausta vastaavalle funktionaalille $\Xi(f)$ saadaan pisteen f_0 ympäristössä

$$\Xi(f) = \|g - (F(f_0) + F'(f_0)(f - f_0))\|^2 + \alpha \|L(f - f_*)\|^2 . \quad (3.24)$$

Rekursiota varten sijoitetaan

$$\begin{aligned} f_0 &\leftarrow f^{(j)} \text{ ja} \\ f &\leftarrow f^{(j+1)} , \end{aligned}$$

jolloin

$$\Xi(f^{(j+1)}) = \|g - F(f^{(j)}) - F'(f^{(j)})(f^{(j+1)} - f^{(j)})\|^2 + \alpha \|L(f^{(j+1)} - f_*)\|^2 . \quad (3.25)$$

Sijoitetaan vielä

$$\begin{aligned} \delta f^{(j)} &\leftarrow f^{(j+1)} - f^{(j)} \text{ ja} \\ g^{(j)} &\leftarrow F(f^{(j)}) , \end{aligned}$$

jolloin linearisoitu funktionaali on

$$\Xi(f) = \|g - g^{(j)} - F'(f^{(j)})\delta f^{(j)}\|^2 + \alpha \|L(\delta f^{(j)} - (f_* - f^{(j)}))\|^2 .$$

Tämän perusmuodossa olevan variaatio-ongelman ratkaisu (minimi) on haettu päivitys $\delta f^{(j)}$. Variaatio-ongelman ratkaisemiseksi kirjoitetaan ongelma LS-muotoon

$$V^{(j)} = \begin{pmatrix} F'(f^{(j)}) \\ \sqrt{\alpha}L \end{pmatrix} \delta f^{(j)} - \begin{pmatrix} g - g^{(j)} \\ \sqrt{\alpha}L(f_* - f^{(j)}) \end{pmatrix} ,$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 72 / 108

Takaisin

Full Screen

Sulje

Poistu



jolloin $\Xi(f^{(j+1)}) = (V^{(j)})^T V^{(j)} = \|V^j\|^2$.

Jos `F.m` on suoran ongelman ratkaiseva Matlab-ohjelma ja `dF.m` laskee kyseisen Jacobin matriisiin F' , on Gauss-Newton -menetelmän ratkaiseva koodifragmentti muotoa

```
f=f0;
for k=1:Niter
    gk=F(f);
    df=[dF(f);sqrt(alpha)*L]\[g-gk;sqrt(alpha)*L*(f_ast-f)];
    f=f+df;
end
```

Luonnollisesti for-silmukan sijasta algoritmi pitää varustaa sopivalla lopetusehdolla, joka usein on yhdistelmä normeista $\|\delta f^{(j)}\|$ ja $\|g - g^{(j)}\|$. Periaatteessa $\delta f^{(j)} \rightarrow 0$, jolloin siis kyseisen lisätyn LS-ongelman residuaali on ortogonaalinen kerroinmatriisin sarakkeiden kanssa (pisteessä $f^{(\infty)}$), eli

$$\begin{pmatrix} F'(f^{(\infty)}) \\ \sqrt{\alpha}L \end{pmatrix}^T \begin{pmatrix} g - g^{(\infty)} \\ \sqrt{(\alpha)}L(f_* - f^{(\infty)}) \end{pmatrix} = 0.$$

Fréchet -derivoituvat epälineaariset sivurajoitteet käsitellään vastaavasti.

Esimerkki 3.4 *Olkoon $g = F(f)$, $F : \mathbb{R}^N \mapsto \mathbb{R}^N$ siten, että*

$$F(f) = Kf + B(f_1^2, \dots, f_N^2)^T \doteq Kf + Bf^2,$$

missä $f \in \mathbb{R}^N$ ja $K, B \in \mathbb{R}^{N \times N}$. Määritään kuvauksen F Jacobin matriisi F' ja regularisoitu GN-iteraatio.

Kirjoitetaan aluksi $F(f)$ muotoon

$$F(f) = \sum_{k=1}^N K_k f_k + \sum_{k=1}^N B_k f_k^2,$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu

◀ ▶

◀ ▶

Sivu 73 / 108

Takaisin

Full Screen

Sulje

Poistu



missä K_k ja B_k ovat matriisien sarakkeet. Tästä muodosta saadaan helposti

$$\frac{\partial F}{\partial f_\ell} = K_\ell + 2B_\ell f_\ell \quad \text{ja}$$

$$\frac{\partial F}{\partial f} = K + 2B \text{diag}(f) ,$$

joten

$$F(f) \approx F(f_0) + F'(f_0)(f - f_0)$$

$$= K f_0 + B f_0^2 + (K + 2B \text{diag}(f_0))(f - f_0) .$$

Siten saadaan rekursio

$$\delta f^{(j)} = \text{sol} \min_{\delta f^{(j)}} \left\| \begin{pmatrix} K + 2B \text{diag}(f^{(j)}) \\ \sqrt{\alpha} L \end{pmatrix} \delta f^{(j)} - \begin{pmatrix} g - F(f^{(j)}) \\ \sqrt{\alpha} L(f_* - f^{(j)}) \end{pmatrix} \right\| \quad \text{ja}$$

$$f^{(j+1)} = f^{(j)} + \text{step} \delta f^{(j)} ,$$

missä *step* tarkoittaa askelpituutta (yleensä < 1).

3.8. Regularisointiparametrin valinta

Perinteisen inversioteorian mukaan regularisointimenetelmä pitää periaatteessa aina varustaa jollakin objektiivisella regularisointiparametrin valintasäännöllä. Valintamenetelmiä, joissa α valitaan ainoastaan operaattorin K ominaisuuksien perusteella, kutsutaan *a priori* -säännöiksi. Menetelmiä, joissa α valitaan perustuen residuaaliin $K f_\alpha - g$, kutsutaan *a posteriori* -menetelmiksi. Käytännössä valintasääntöjen ohella käytetään aina myös visuaalista arviointia, toisin sanoen lähes aina yksi *a posteriori* -kriteeri on ratkaisujen visuaalinen luonne. Yksi mahdollisuus on myös käyttää simulointeja regularisointiparametrin valinnan apuna, mikäli ratkaisujen *luokka* ja virheiden ominaisuudet tunnetaan riittävän hyvin.

Regularisointiparametrien valintamenetelmiä ovat muun muassa *Morozovin diskrepanssiperiaate*, *L-käyrä -menetelmä*, *yleistetty ristiarviointi* (*generalized cross validation*) ja *kvasioptimaalisuus-kriteeri*. Tässä käsitellään vain kahta ensin mainittua menetelmää.

Sisältö:

1. Johdanto
2. Yleistetty käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 74 / 108

Takaisin

Full Screen

Sulje

Poistu



3.8.1. Morozovin diskrepanssiperiaate

Tarkastellaan tavallista Tihonov-regularisointia. Olkoon $K : \mathcal{H}_1 \mapsto \mathcal{H}_2$. Jos $\mathcal{R}(K) = \mathcal{H}_2$, löytyy periaatteessa aina ratkaisu siten, että kohinaisellekin datalle $\|Kf_\alpha^\delta - g^\delta\| = 0$, mutta kuten aikaisemmin on todettu, minimoiva ratkaisu $f = K^\dagger g^\delta$ on tällöin tavallisesti merkityksetön. Morozovin diskrepanssiperiaatteen mukaisesti ”ei kannata yrittää” saavuttaa pienempää virheen normia kuin tämän voidaan olettaa olevan. Siis, jos on olemassa arvio

$$\|g - g^\delta\| = \delta ,$$

tulee regularisointiparametri valita siten, että

$$\|Kf_\alpha^\delta - g^\delta\| = \delta . \quad (3.26)$$

Voidaan osoittaa, että $\alpha = \alpha(\delta)$, jolle (3.26) toteutuu, on olemassa, jos $\|g^\delta\| \geq \delta$ ja häiriöttömälle datalle pätee $g \in \mathcal{R}(K)$. Tämä nähdään seuraavasti. Kuten aikaisemminkin, kirjoitetaan

$$g^\delta = \underbrace{\sum_{k=1}^{\infty} \langle g^\delta, u_k \rangle u_k}_{P_{\mathcal{R}(K)} g^\delta} + P_{\mathcal{R}(K)^\perp} g^\delta ,$$

$$f_\alpha^\delta = \sum_{k=1}^{\infty} \frac{\sigma_k}{\sigma_k^2 + \alpha} \langle g^\delta, u_k \rangle v_k \quad , \quad K v_k = \sigma_k u_k ,$$

$$K f_\alpha^\delta = \underbrace{\sum_{k=1}^{\infty} \frac{\sigma_k^2}{\sigma_k^2 + \alpha} \langle g^\delta, u_k \rangle u_k}_{\in \mathcal{R}(K)} \quad \text{ja}$$

$$K f_\alpha^\delta - g^\delta = \sum_{k=1}^{\infty} \frac{-\alpha}{\sigma_k^2 + \alpha} \langle g^\delta, u_k \rangle u_k - P_{\mathcal{R}(K)^\perp} g^\delta ,$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 75 / 108

Takaisin

Full Screen

Sulje

Poistu



missä $P_{\mathcal{R}(K)^\perp}$ on projektori $P : \mathcal{H}_2 \mapsto \mathcal{R}(K)^\perp$, jolloin siis $P_{\mathcal{R}(K)^\perp}g = 0$, koska $g \in \mathcal{R}(K)$. Siten ratkaisun virheelle saadaan

$$\|Kf_\alpha^\delta - g^\delta\|^2 = \sum_{k=1}^{\infty} \left(\frac{-\alpha}{\sigma_k^2 + \alpha} \right)^2 |\langle g^\delta, u_k \rangle|^2 + \|P_{\mathcal{R}(K)^\perp}g^\delta\|^2.$$

Nyt kuvaus

$$F(\alpha) = \|Kf_\alpha^\delta - g^\delta\|$$

on jatkuva ja kasvava funktio, jolle saadaan raja-arvot

$$\begin{aligned} \lim_{\alpha \rightarrow 0^+} F(\alpha) &= \|P_{\mathcal{R}(K)^\perp}g^\delta\| = \|P_{\mathcal{R}(K)^\perp}g^\delta - \underbrace{P_{\mathcal{R}(K)^\perp}g}_{=0}\| \\ &= \|P_{\mathcal{R}(K)^\perp}(g^\delta - g)\| \leq \|g^\delta - g\| = \delta \quad \text{ja} \\ \lim_{\alpha \rightarrow \infty} F(\alpha) &= \|g^\delta\| > \delta. \end{aligned}$$

Väitetty tulos seuraa suoraan [Bolzanon](#) lauseen perusteella. Voidaan osoittaa, että Morozovin diskrepanssiperiaatteen perusteella tapahtuva regularisointiparametrin valinta johtaa säännölliseen algoritmiin.

3.8.2. L-käyrä -menetelmä

L-käyrä on graafi

$$\alpha \mapsto (\log \|Kf_\alpha^\delta - g^\delta\|, \log \|Lf_\alpha^\delta\|),$$

jonka muoto on yleensä/usein/joskus/sattumalta kirjaimen “L” muotoinen. L-käyrä -menetelmän mukaisesti valitaan regularisointiparametri siten, että kyseisellä graafilla “ollaan hieman kulumasta oikealla”. Tässä ei esitetä tälle valintamenetelmälle teoreettisia perusteita. Menetelmä on säännöllinen algoritmi vain “suppealle” operaattoriluokalle ja sen perustelut ovat muissa tapauksissa kvalitatiivisia ja heuristisia. Tämä ei taas tarkoita, etteikö menetelmää voisi käyttää, jos graafi on kuvatun kaltainen.

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 76 / 108

Takaisin

Full Screen

Sulje

Poistu



Esimerkki 3.5 Tarkastellaan johdantoluvun derivaattaesimerkin regularisointiparametrin valintaa. Asetetaan $\delta^2 = N^{-1}E \{e^T e\} = 0.1$ ja käytetään regularisointimatriisia $L = I$. Koska kyse on simulaatiosta, voidaan tarkastella myös todellista estimointivirhettä $\|f - f_\alpha^\delta\|$.

Kuvassa 3.1 on esitetty todellisen estimointivirheen $\|f - f_\alpha^\delta\|$ riippuvuus regularisointiparametrista α kolmella eri kohinatasolla δ . Optimaalisen regularisointiparametrin riippuvuus kohinatasosta on ilmeinen. Kuvassa 3.2 taas on esitetty sivurajoitteen normin $\|L f_\alpha^\delta\|$ riippuvuus päärajoitteen normista $\|K f_\alpha^\delta - g^\delta\|$ samoilla kohinatasoilla. L -käyrän käyttökelpoisuus riippuu selvästi kohinatasosta. Kuvassa 3.3 on esitetty regularisointiparametrin määrääminen Morozovin periaatteen ja L -käyrän avulla, kun $\delta = 0.1$.

Näiden kuvien kuvaajat ovat siis graafit

$$\begin{aligned} \text{Morozov} & (\alpha, \|K f_\alpha^\delta - g^\delta\|) \quad \text{ja} \\ L\text{-käyrä} & (\log \|K f_\alpha^\delta - g^\delta\|, \log \|L f_\alpha^\delta\|), \end{aligned}$$

missä $\alpha \in [10^{-8}, 10^{-1}]$. \square

3.9. Tihonov-regularisointi ja normaali jakaumaoletukset

Tässä kappaleessa osoitetaan, että Tihonov-regularisointi vastaa tiettyjä erittäin rajoitettuja oletuksia sekä havaintovirheistä että itse estimoitavasta suureesta. Jatkossa sekä satunnaismuuttujia X että näiden saamia arvoja x merkitään pienillä kirjaimilla, jos sekaantumisen vaaraa ei ole

3.9.1. Mean square -estimointi

Tarkastellaan äärellisdimensioisia tapauksia, jolloin $f \in \mathbb{R}^N$ ja $g \in \mathbb{R}^M$. Tilastollisessa inversio-teoriassa tulkitaan sekä muuttuja f että havaintovirhe ν satunnaismuuttujiksi (jolloin havainto g on myös satunnaismuuttuja). Tarkastellaan tilastollista minimointiongelmää (*minimum mean square error*, MMSE)

$$\hat{f} = \text{sol} \min_f E \left\{ \|f - \hat{f}\|^2 \right\}, \quad (3.27)$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 77 / 108

Takaisin

Full Screen

Sulje

Poistu



Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



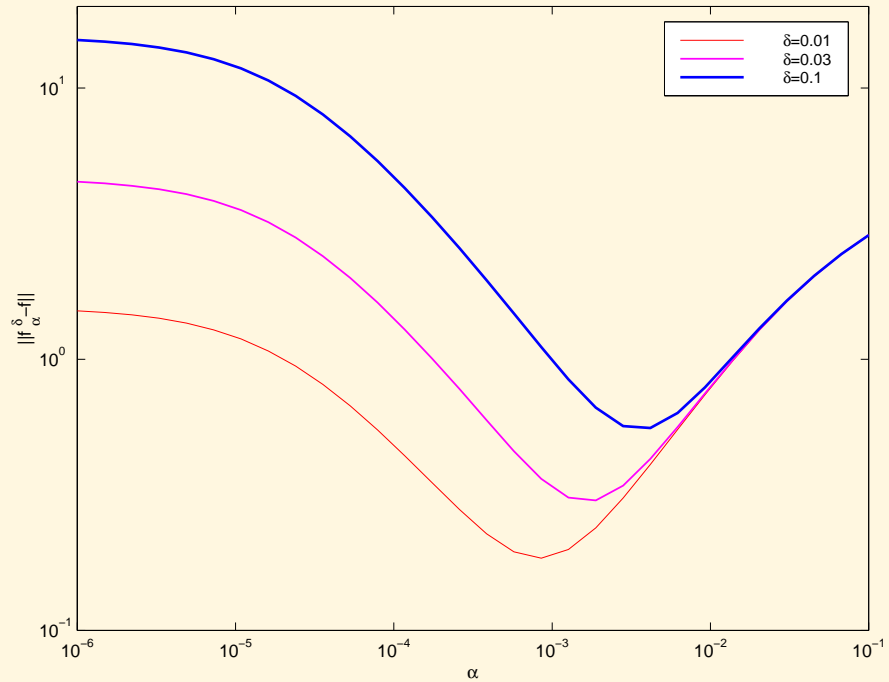
Sivu 78 / 108

Takaisin

Full Screen

Sulje

Poistu



Kuva 3.1: Estimointivirheen $\|f - f_\alpha^\delta\|$ riippuvuus regularisointiparametrasta α kohinatasoilla $\delta = 0.01$ (ohuin viiva), $\delta = 0.03$ ja $\delta = 0.1$ (paksuin viiva). [\[KOODI\]](#)



Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



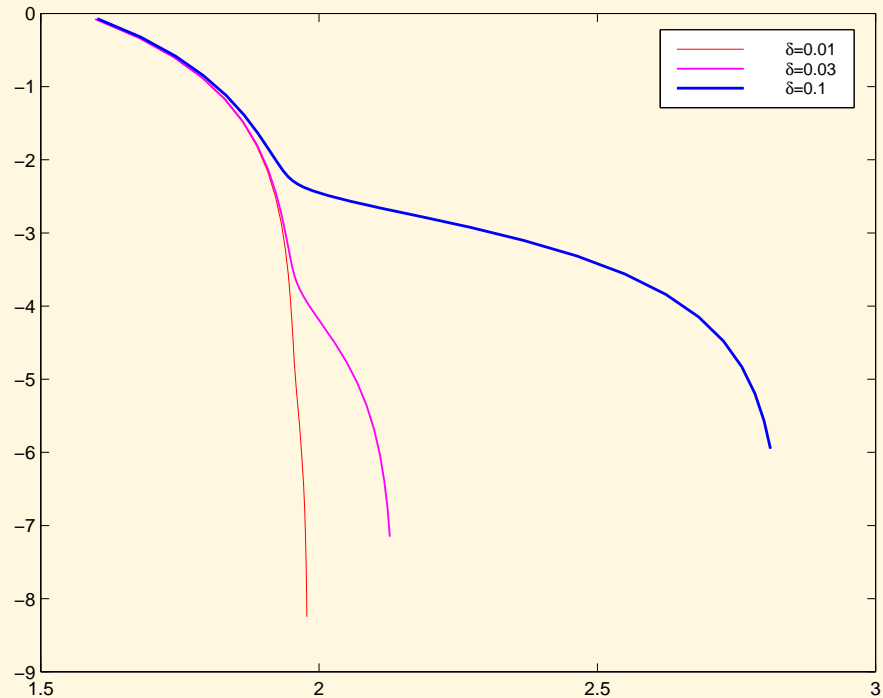
Sivu 79 / 108

Takaisin

Full Screen

Sulje

Poistu

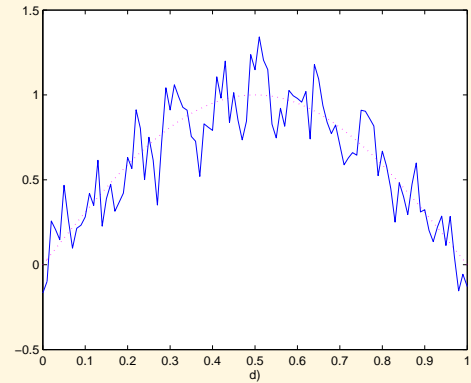
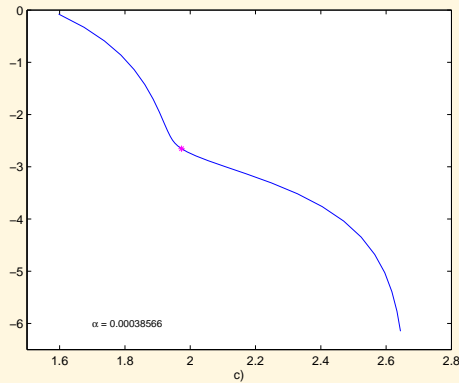
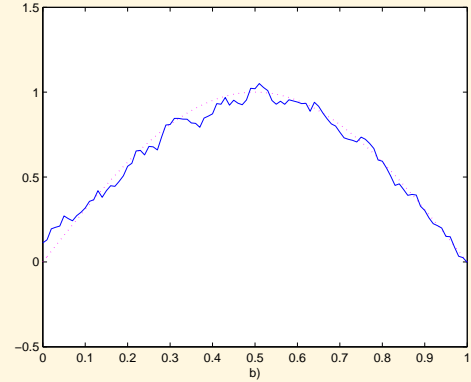
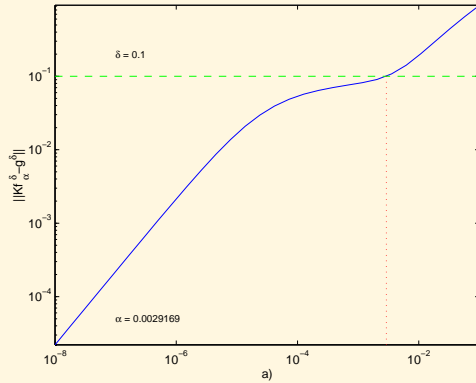


Kuva 3.2: Sivurajoitteen normin $\|Lf_\alpha^\delta\|$ riippuvuus päärajoitteen normista $\|Kf_\alpha^\delta - g^\delta\|$ kohinata-soilla $\delta = 0.01$ (ohuin viiva), $\delta = 0.03$ ja $\delta = 0.1$ (paksuin viiva). Molempien akselien numerointi viittaa kymmenen eksponentteihin. [\[KOODI\]](#)



Sisältö:

1. Johdanto
2. Yleistetyt käänneiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä



Kuva 3.3: a) Morozovin diskrepanssimenetelmä ja b) tätä vastaava estimaatti ($\alpha = 0.002$, $\delta = 0.1$) sekä c) L-käyrä -menetelmä (logaritmiset akselit) ja d) tätä vastaava estimaatti ($\alpha = 0.0002$, $\delta = 0.1$). [\[KOODI\]](#)

Etusivu



Sivu 80 / 108

Takaisin

Full Screen

Sulje

Poistu



missä E on odotusarvo-operaattori ja satunnaismuuttujan f tiheysfunktio on $p_f(f)$. Merkitään $f \sim p_f$. Jos $g = Kf$ ja havaintovirheet $\nu \sim p_\nu$ ovat additiivisia, voidaan havaintoyhtälöksi kirjoittaa

$$\begin{aligned} g^\delta &= g + \nu, \\ Kf &= g = g^\delta - \nu \quad \text{ja} \\ g^\delta - Kf &= \nu. \end{aligned}$$

Oletetaan jatkossa, että f ja ν ovat toisistaan riippumattomia. Tilanne voidaan tulkita seuraavasti: tiheysfunktioista p_f “arvotaan” alkio (parametri) f_k . Tämän jälkeen “lasketaan” kuva $g_k = Kf_k$, joka on nyt myös satunnaismuuttuja. Tämän jälkeen arvotaan satunnaismuuttuja (kohina, häiriö, havaintovirhe) ν_k ja “lisätään” tämä kuvaan g_k . Useissa tapauksissa voidaan satunnaismuuttujan ν katsoa olevan (tilastollisesti) riippumaton satunnaismuuttujasta f . Vektori $g_k^\delta = Kf_k + \nu_k$ on todellinen havainto, kuva 3.4.

Minimimiestimaatin määräämisongelmaa voidaan tulkita kokeellisesti seuraavasti.

1. Generoidaan N vektoria $f_k \in \mathcal{H}_1$ jakaumasta p_f .
2. Lasketaan (virheettömät/kohinattomat) havainnot $g_k = Kf_k$.
3. Generoidaan N virhevektoria $\nu_k \in \mathcal{H}_2$ jakaumasta p_ν .
4. Lasketaan kohinaiset havainnot $g_k^\delta = g_k + \nu_k$.
5. Määrätään $\hat{f}_k = F(g_k^\delta)$ siten, että virheet $f_k - \hat{f}_k$ ovat “keskimäärin” mahdollisimman pieniä koko joukolla $\{f_k, k = 1, \dots, N\}$, eli

$$\sum_{k=1}^N \|f_k - \hat{f}_k\|^2 \longrightarrow \min .$$

Miten kuvaus F määrätään?

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 81 / 108

Takaisin

Full Screen

Sulje

Poistu



Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



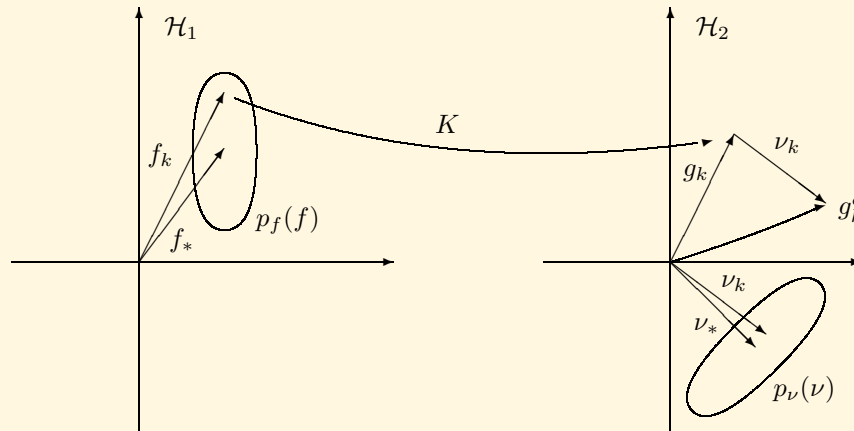
Sivu 82 / 108

Takaisin

Full Screen

Sulje

Poistu



Kuva 3.4: Havainnon g_k^δ määräytyminen, kun f_k ja ν_k ovat satunnaismuuttujia.



Merkintöjen yksinkertaistamiseksi kirjoitetaan jatkossa $g = g^\delta = Kf + \nu$. Kriteerin (3.27) minimoiva estimaatti \hat{f} on ehdollinen odotusarvo (todistus)

$$\hat{f} = E\{f|g\} = \int f p_{f|g}(f) df,$$

missä $p_{f|g}$ on satunnaismuuttujan f ehdollinen jakauma tunnetulla g . Bayesin kaavan mukaan

$$p_{f|g} p_g = p_{g|f} p_f.$$

Jakaumaa $p_{g|f}$ kutsutaan uskottavuudeksi (*likelihood*) ja se voidaan tulkita havainnon g todennäköisyydeksi, kun f on tunnettu. Jakaumaa p_f kutsutaan priorijakaumaksi (*a priori*) ja se taas tulkitaan todennäköisyydeksi, että tuntematon muuttuja saa arvon f . Jakaumaa $p_{f|g}$ kutsutaan posteriorijakaumaksi, jonka määrittäminen on tilastollisen inversioteorian ensisijainen tavoite. Jakauman perusteella voidaan vastata kaikkiin kysymyksiin koskien satunnaismuuttujaa f . Voidaan osoittaa, että jos f ja ν ovat toisistaan riippumattomia, on $\nu = g - Kf \sim p_\nu$, joten (todistus)

$$p_{g|f}(g) = p_\nu(g - Kf).$$

Siten

$$p_{f|g} = p_g^{-1} p_\nu(g - Kf) p_f(f), \quad (3.28)$$

missä p_g ei riipu satunnaismuuttujasta f .

3.9.2. Normaalijakauma

Tarkastellaan nyt erikoistapausta, jossa sekä $f \in \mathbb{R}^N$ että $\nu \in \mathbb{R}^M$ ovat riippumattomia ja normaalijakautuneita siten, että

$$p_f(f) = \frac{1}{(2\pi)^{N/2} |\Gamma_f|^{1/2}} \exp\left(-\frac{1}{2}(f - f_*)^T \Gamma_f^{-1} (f - f_*)\right) \quad \text{ja}$$

$$p_\nu(\nu) = \frac{1}{(2\pi)^{M/2} |\Gamma_\nu|^{1/2}} \exp\left(-\frac{1}{2}(\nu - \nu_*)^T \Gamma_\nu^{-1} (\nu - \nu_*)\right),$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 83 / 108

Takaisin

Full Screen

Sulje

Poistu



missä f_* ja ν_* ovat satunnaismuuttujien f ja ν odotusarvot sekä Γ_f ja Γ_ν kovarianssimatriisit. Silloin myös Kf ja $Kf + \nu$ ovat normaalijakautuneita. Posteriorijakauma on siten muotoa

$$\begin{aligned} p_{f|g} &\propto p_{g|f} p_f = p_\nu(g - Kf) p_f(f) \\ &\propto \exp\left(-\frac{1}{2}(g - Kf - \nu_*)^T \Gamma_\nu^{-1}(g - Kf - \nu_*)\right) \\ &\quad \exp\left(-\frac{1}{2}(f - f_*)^T \Gamma_f^{-1}(f - f_*)\right) \\ &= \exp\left(-\frac{1}{2}Q(f)\right), \end{aligned}$$

missä $Q(f)$ on neliöllinen funktio

$$Q(f) = (g - Kf - \nu_*)^T \Gamma_\nu^{-1}(g - Kf - \nu_*) + (f - f_*)^T \Gamma_f^{-1}(f - f_*). \quad (3.29)$$

Normaalijakauman symmetriaominaisuuksien vuoksi posteriorijakauman odotusarvo on jakauman maksimikohta, joten

$$\begin{aligned} \hat{f} &= \operatorname{sol} \max_f p_{f|g} \\ &= \operatorname{sol} \min_f Q(f). \end{aligned}$$

Olkoot L_ν ja L_f matriisien Γ_ν^{-1} ja Γ_f^{-1} Cholesky-tekijät (tai muut matriisineliojuuret) eli $L_f^T L_f = \Gamma_f^{-1}$ ja L_ν vastaavasti. Minimoitava funktio $Q(f)$ voidaan kirjoittaa muotoon

$$\begin{aligned} \hat{f} &= \operatorname{sol} \min_f \left\{ (L_\nu(Kf - g - \nu_*))^T (L_\nu(Kf - g - \nu_*)) \right. \\ &\quad \left. + (L_f(f - f_*))^T (L_f(f - f_*)) \right\} \\ &= \operatorname{sol} \min_f \left\{ \|L_\nu(Kf - g - \nu_*)\|^2 + \|L_f(f - f_*)\|^2 \right\} \end{aligned} \quad (3.30)$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 84 / 108

Takaisin

Full Screen

Sulje

Poistu



$$= \operatorname{sol} \min_f \left\| \begin{pmatrix} L_\nu K \\ L_f \end{pmatrix} f - \begin{pmatrix} L_\nu (g - \nu_*) \\ L_f f_* \end{pmatrix} \right\|^2. \quad (3.31)$$

Tämän formaali ratkaisu on (laskuharjoitukset)

$$\hat{f} = \left(K^T \Gamma_\nu^{-1} K + \Gamma_f^{-1} \right)^{-1} \left(K^T \Gamma_\nu^{-1} (g - \nu_*) + \Gamma_f^{-1} f_* \right). \quad (3.32)$$

Kriteeri (3.31) on selvästi samaa muotoa kuin yleistetty Tihonov-regularisointi. Siten (yleistetty) Tihonov-regularisointi vastaa muodollisesti pienimmän neliövirheen estimointia normaali-jakaumaoletuksilla! Erityisesti Tihonov-regularisoinnissa tehdään esimerkiksi tapauksessa $L = I$ implisiittisesti seuraavat oletukset:

- Ratkaisuvektorin f alkioit ovat keskenään tilastollisesti riippumattomia keskiarvolla f_* ja lisäksi vektorin f alkioiden varianssit ovat yhtä suuret.
- Havaintovirhevektorin alkioit ovat myös keskenään tilastollisesti riippumattomia keskiarvolla $\nu_* = 0$ ja lisäksi virheiden ν_k varianssit ovat yhtä suuret (virhe on “jakautunut tasaisesti” havaintoihin).

Mikäli parametrin f jakaumalle voidaan tehdä normaalijakauma-aproksimaatio (odotusarvo f_* ja kovarianssi Γ_f), saadaan neliövirhekriteerin mielessä “oikeat” matriisit kuten yllä on esitetty.

Esimerkki 3.6 Havainnollistetaan kovarianssien Γ_ν ja Γ_f merkitystä seuraavalla esimerkillä. Olkoon $K : \mathbb{R}^2 \mapsto \mathbb{R}^2$ siten, että

$$K = \begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix},$$

jolloin $\dim \mathcal{R}(K) = 1$ sekä

$$\mathcal{R}(K) = \{c_1(2, 1)^T, c_1 \in \mathbb{R}\} \quad \text{ja} \quad \mathcal{N}(K) = \{c_2(-2, 1)^T, c_2 \in \mathbb{R}\}.$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 85 / 108

Takaisin

Full Screen

Sulje

Poistu



Oletetaan lisäksi, että $f = (3, 0)^T$, jolloin $Kf = (6, 3)^T$. Lisätään vektoriin Kf kohinaa ν , ja olkoon havainto siten $g = Kf + \nu$. Olkoot f ja ν normaalijakautuneita siten, että niiden odotusarvot ovat $f_* = \nu_* = (0, 0)^T$ ja kovarianssimatriisit

$$\Gamma_\nu = \begin{pmatrix} 10 & -1 \\ -1 & 2 \end{pmatrix} \quad \text{ja} \quad \Gamma_f = \begin{pmatrix} 10 & 0 \\ 0 & 1 \end{pmatrix}.$$

Määritään g :n keskihajontaellipsi ja lasketaan tämän ellipsin pisteitä vastaten

- Minimimormiratkaisut $f_1 = K^\dagger$ sekä
- näitä vastaavat (ortogonaaliset) projektiot $g_1 = Kf_1$ maaliavaruudelle $\mathcal{R}(K)$.
- Pienimmän (tilastollisen) neliövirheen estimaatit $f_{\text{MMSE}} = (K^T \Gamma_\nu^{-1} K + \Gamma_f^{-1})^{-1} K^T \Gamma_n u^{-1} g$ sekä
- näitä vastaavat (ei-ortogonaaliset) projektiot $g_{\text{MMSE}} = Kf_{\text{MMSE}}$ maaliavaruudelle $\mathcal{R}(K)$.

Laskennan tulokset on esitetty kuvassa 3.5. Havaitaan, että valittu $f = (3, 0)^T$ on "todennäköinen" näyte jakaumasta $\mathcal{N}(0, \Gamma_f)$, joten MMSE-ratkaisut ovat selvästi parempia kuin miniminormi-ratkaisut. Siten oikeiden kovarianssien käyttäminen on (tämänkin esimerkin valossa) perusteltua. Toisaalta, jos jakauma $\mathcal{N}(0, \Gamma_f)$ olisi väärä, saataisiin aina estimaatteja f_{MMSE} , jotka silti näyttäisivät olevan tästä jakaumasta. Jos esimerkiksi $f = (0, 1.5)^T$, olisi edelleen $Kf = (6, 3)^T$ ja kaikki estimaatit olisivat kuten tapauksessa $f = (3, 0)^T$. Nyt kuitenkin $p_f(3, 0) = 1.96p_f(0, 1.5)$, joten estimaattori ajaa ratkaisuja (priorijakauman kannalta) todennäköisempiä havaintoja kohti. Jos f_* ja ν_* eivät olisi nollia, olisi tulosten ero tässä esitettyä huomattavampi. \square

Voidaan sanoa, että väärät kovarianssivalinnat voivat aiheuttaa suuria virheitä satunnaismuuttujan f estimoinnissa. Toisaalta kuitenkin myös tavallisen Tihonov-regularisoinnin implisiittiset valinnat $f_* = 0$, $\nu_* = 0$, $\Gamma_f = \alpha^{-1}I$ ja $\Gamma_\nu = I$ voivat olla tilanteeseen sopimattomia.

Tehtäviä

Sisältö:

1. Johdanto
2. Yleistetyt käännekuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 86 / 108

Takaisin

Full Screen

Sulje

Poistu



Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



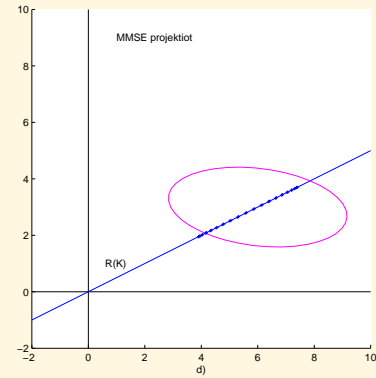
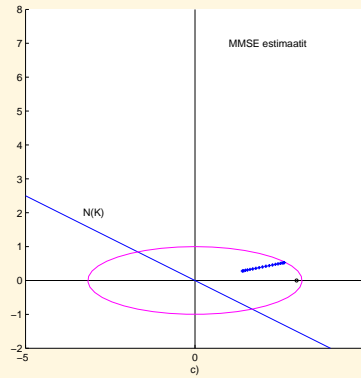
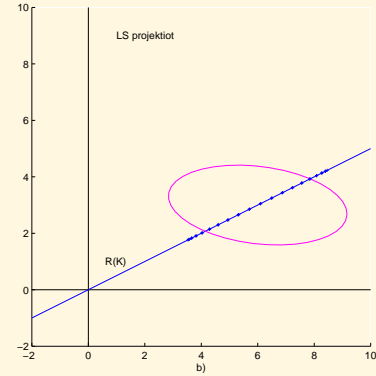
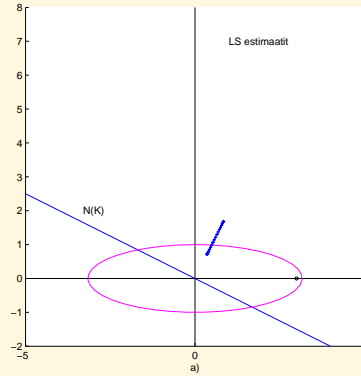
Sivu 87 / 108

Takaisin

Full Screen

Sulje

Poistu



Kuva 3.5: a) ja c) Jakauman p_f keskihajontaellipsit, nolla-avaruus $\mathcal{N}(K)$ ja “oikea” f (ympyrä) sekä a) miniminormiestimaatit $f_1 = K^\dagger g$ (+), c) Pienimmän (tilastollisen) neliövirheen estimaatit $f_{\text{MMSE}} = (K^T \Gamma_\nu^{-1} K + \Gamma_f^{-1})^{-1} K^T \Gamma_\nu^{-1} g$ (+). b) ja d) Jakauman $Kf + p_\nu$ keskihajontaellipsit ja maalijoukko $\mathcal{R}(K)$ sekä b) ortogonaaliset projektiot $g_1 = K f_1$ (+), d) projektiot $g_{\text{MMSE}} = K f_{\text{MMSE}}$ (+). Huomaa, kuinka kuvassa a) selvästi $K^\dagger g \in \mathcal{N}(K)^\perp$. **[KOODI]**

**Sisältö:**

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 88 / 108

Takaisin

Full Screen

Sulje

Poistu

Joissakin tapauksissa ongelman $g = Kf$ matriisin K dimensiot ovat niin suuret, että suorat menetelmät ongelman ratkaisemiseksi ovat liian raskaita. Tällöin on mahdollista pyrkiä saamaan ratkaisu iteratiivisesti. Varsinkin, jos matriisi K on harva, osoittautuvat tietyt iteratiiviset menetelmät erittäin tehokkaiksi.

Perinteisesti käänteisongelmien kannalta oleellista on lisäksi, että tietyt iteratiiviset algoritmit varustettuna sopivalla iteraatioiden terminointisäännöllä osoittautuvat säännöllisiksi algoritmeiksi. Tällaisia algoritmeja ovat mm. Landweber-, konjugaattigradientti- ja Kaczmarz-iteraatiot.

4.1. Kiintopisteiteraatiot

Olkoon $\mathcal{S} \subset \mathcal{H}$ suljettu ja $T : \mathcal{S} \mapsto \mathcal{S}$. Operaattori T on *kutistava kuvaus* joukossa \mathcal{S} , jos on olemassa vakio $0 \leq \kappa < 1$ siten, että $\|T(f_1) - T(f_2)\| \leq \kappa \|f_1 - f_2\|$ kaikille $f_1, f_2 \in \mathcal{S}$.

Jos \mathcal{S} on konvekksi ja $\|T'(f)\| \leq \kappa < 1$ kaikille $f \in \mathcal{S}$, on T kutistava, koska väliarvolauseen mukaan $\|T(f_1) - T(f_2)\| \leq \sup \|T'(f)\| \|f_1 - f_2\|$, $f \in \mathcal{S}$.



Olkoon $f = T(f)$ jollekin $f \in \mathcal{H}$. Tällaista pistettä kutsutaan kuvauksen T kiintopisteeksi. Muotoa $f_{k+1} = T(f_k)$ olevaa peräkkäisten approksimaatioiden jonoa kutsutaan kiintopisteiteraatioksi, mikäli se suppenee johonkin raja-arvoon f_* .

Tarkastellaan Newtonin menetelmää ja merkitään $x = T(x)$ ja $f(x) = 0$. Tällöin

$$T(x) = x - \frac{f(x)}{f'(x)} = x,$$

joten

$$\frac{f(x)}{f'(x)} = 0, \quad \text{jos } f'(x) \neq 0.$$

Näin ollen

$$T'(x) = \frac{f(x)f''(x)}{(f'(x))^2} \Big|_{x_{pp}} = 0,$$

joten kyseessä on kvadraattinen konvergenssi.

Lause 4.1 KIINTOPISTELAUSE. *Olkoon \mathcal{H} täydellinen ja $T : \mathcal{H} \mapsto \mathcal{H}$ kutistava kuvaus suljetussa joukossa $\mathcal{S} \subset \mathcal{H}$. Silloin on olemassa yksikäsitteinen vektori $f_* \in \mathcal{S}$, jolle $f_* = T(f_*)$. Lisäksi tämä vektori saadaan kiintopisteiteraatiolla lähtien mielivaltaisesta vektorista $f_1 \in \mathcal{S}$.*

TODISTUS. Olkoon $f_1 \in \mathcal{S}$ ja $f_{k+1} = T(f_k)$. Tällöin

$$\|f_{k+1} - f_k\| = \|T(f_k) - T(f_{k-1})\| \leq \kappa \|f_k - f_{k-1}\|.$$

Siten

$$\|f_{k+1} - f_k\| \leq \kappa^{k-1} \|f_2 - f_1\|.$$

Edelleen saadaan

$$\begin{aligned} \|f_{k+p} - f_k\| &\leq \|f_{k+p} - f_{k+p-1}\| + \|f_{k+p-1} - f_{k+p-2}\| + \cdots + \|f_{k+1} - f_k\| \\ &\leq (\kappa^{k+p-2} + \kappa^{k+p-3} + \cdots + \kappa^{k-1}) \|f_2 - f_1\| \end{aligned}$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 89 / 108

Takaisin

Full Screen

Sulje

Poistu



$$\begin{aligned} &\leq \left(\kappa^{k-1} \sum_{j=0}^{\infty} \kappa^j \right) \|f_2 - f_1\| \\ &= \frac{\kappa^{k-1}}{1 - \kappa} \|f_2 - f_1\|. \end{aligned}$$

Siten $\{f_k\}$ on Cauchy-jono. Koska \mathcal{H} on täydellinen ja \mathcal{S} suljettu, jonolla on raja-arvo $f_k \rightarrow f_* \in \mathcal{S}$. Osoitetaan, että $f_* = T(f_*)$. Koska

$$\begin{aligned} \|f_* - T(f_*)\| &= \|f_* - f_k + f_k - T(f_*)\| \\ &\leq \|f_* - f_k\| + \|f_k - T(f_*)\| \\ &= \|f_* - f_k\| + \|T(f_{k-1}) - T(f_*)\| \\ &\leq \|f_* - f_k\| + \kappa \|f_{k-1} - f_*\|, \end{aligned}$$

saadaan viimeinen lauseke mielivaltaisen pieneksi valitsemalla k riittävän suureksi. Siten $\|f_* - T(f_*)\| = 0$ ja $f_* = T(f_*)$. Olkoot lopuksi f_* ja f_+ kaksi kiintopistettä. Silloin

$$\|f_* - f_+\| = \|T(f_*) - T(f_+)\| \leq \kappa \|f_* - f_+\|.$$

Koska $\kappa < 1$, on oltava $f_* - f_+ = 0$, joten kiintopiste on yksikäsitteinen. \square

Tarkastellaan seuraavaksi lineaarista yhtälöä

$$g = Kf, \tag{4.1}$$

missä $K : \mathcal{H} \mapsto \mathcal{H}$. Yhtälö (4.1) voidaan kirjoittaa muotoon

$$f = (I - K)f + g \doteq T(f).$$

Nyt $T'(f) \equiv I - K$, joten yhtälö voidaan ratkaista kiintopisteiteraatiolla, mikäli $\|I - K\| < 1$.

Jos $\|I - K\| \geq 1$ voidaan merkitä $\tilde{g} = cg$ ja $\tilde{K} = cK$ ja valitaan $c \in \mathbb{R}$ siten, että $\|I - \tilde{K}\| < 1$.

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 90 / 108

Takaisin

Full Screen

Sulje

Poistu



4.2. Landweber-iteraatiot

Miniminnormiratkaisulle $f = K^\dagger g$ pätee $f \in \mathcal{N}(K)^\perp$ ja $K^*Kf = K^*g$. Kun tämä toteutuu, on myös

$$f = f + \beta(K^*g - K^*Kf)$$

kaikille $\beta \in \mathbb{R}$. Pyritään kiintopisteiteraatioon ja asetetaan

$$f_{k+1} = f_k + \beta(K^*g - K^*Kf_k),$$

eli

$$T(f) = f + \beta(K^*g - K^*Kf).$$

Olkoon $f_0 = 0$ (jos $f_0 \neq 0$ voidaan tehdä muuttujan vaihto $f' = f - f_0$). Silloin $f_k \in \mathcal{R}(K^*) = \mathcal{N}(K)^\perp$ kaikille k , sillä

$$\begin{aligned} f_1 &= \underbrace{f_0}_{\in \mathcal{R}(K^*)} + \beta \left(\underbrace{K^*g}_{\in \mathcal{R}(K^*)} + \underbrace{K^*Kf_0}_{\in \mathcal{R}(K^*)} \right), \\ f_2 &= \underbrace{f_1}_{\in \mathcal{R}(K^*)} + \beta \left(\underbrace{K^*g}_{\in \mathcal{R}(K^*)} + \underbrace{K^*Kf_1}_{\in \mathcal{R}(K^*)} \right), \\ &\vdots \end{aligned}$$

Siten, jos $f_k \rightarrow f$ suppenee, niin $f \in \mathcal{N}(K)^\perp$ ja $K^*g = K^*Kf$. Määritellään virhe $r_k = f_k - f$, jolloin

$$\begin{aligned} r_{k+1} &= r_k + \beta(K^*g - K^*K(r_k + f)) \\ &= r_k + \beta(\underbrace{K^*g - K^*Kf}_{=0} - K^*Kr_k) \\ &= (I - \beta K^*K)r_k \quad \text{ja} \\ r_k &= (I - \beta K^*K)^k r_0. \end{aligned} \tag{4.2}$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 91 / 108

Takaisin

Full Screen

Sulje

Poistu



Siten suppenemisen ehtona on $\|I - \beta K^* K\| < 1$. Määrätään suurin β , jolle tämä ehto on voimassa. Olkoon $\{\lambda_\ell, v_\ell\}$ operaattorin $K^* K$ ominaissysteemi. Koska $r_0 = -f \in \mathcal{N}(K)^\perp$, voidaan kirjoittaa

$$r_0 = \sum_{\lambda_\ell > 0} \langle r_0, v_\ell \rangle v_\ell \quad \text{ja} \quad (4.3)$$

$$r_k = (I - \beta K^* K)^k \sum_{\lambda_\ell > 0} \langle r_0, v_\ell \rangle v_\ell \quad (4.4)$$

$$= \sum_{\lambda_\ell > 0} \langle r_0, v_\ell \rangle (I - \beta K^* K)^k v_\ell$$

$$= \sum_{\lambda_\ell > 0} \langle r_0, v_\ell \rangle (I - \beta K^* K)^{k-1} \underbrace{(I - \beta K^* K)}_{(1 - \beta \lambda_\ell) v_\ell} v_\ell$$

$$= \sum_{\lambda_\ell > 0} \langle r_0, v_\ell \rangle (1 - \beta \lambda_\ell)^k v_\ell \quad (\text{induktio}), \quad (4.5)$$

josta virheen normin neliöksi saadaan edelleen

$$\|r_k\|^2 = \sum_{\lambda_\ell > 0} (1 - \beta \lambda_\ell)^{2k} |\langle r_0, v_\ell \rangle|^2. \quad (4.6)$$

Suppeneminen edellyttää, että $\|r_k\| \rightarrow 0$, joten on oltava $|1 - \beta \lambda_\ell| < 1$ kaikille ℓ . Näin ollen

$$\begin{aligned} |1 - \beta \lambda_\ell| < 1 &\Leftrightarrow -1 < 1 - \beta \lambda_\ell < 1 \\ &\Leftrightarrow -2 < -\beta \lambda_\ell < 0 \\ &\Leftrightarrow 0 < \beta \lambda_\ell < 2 \\ &\Leftrightarrow 0 < \beta < \frac{2}{\lambda_\ell}. \end{aligned}$$

Tällöin $|1 - \beta \lambda_\ell|^{2k} \rightarrow 0$ kaikille ℓ ja edelleen $\|r_k\| \rightarrow 0$. Koska $f_k \in \mathcal{N}(K)^\perp$, on oltava $f_k \rightarrow K^\dagger g$, sillä kyseisessä aliavaruudessa vain $K^\dagger g$ toteuttaa ehdon $K^* K f = K^* g$. Algoritmin huonona puolena on, että se suppenee hyvin hitaasti.

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 92 / 108

Takaisin

Full Screen

Sulje

Poistu



Tarkastellaan nyt algoritmin käyttäytymistä kohinaisen datan g^δ tapauksessa, kun $\|g - g^\delta\| \leq \delta$. Olkoon $d_k^\delta = f_k^\delta - f_k$. Siten

$$f_{k+1}^\delta = f_k^\delta + \beta (K^* g^\delta - K^* K f_k^\delta) \quad \text{ja} \quad (4.7)$$

$$\begin{aligned} d_{k+1}^\delta &= f_{k+1}^\delta - f_{k+1} \\ &= f_k^\delta + \beta (K^* g^\delta - K^* K f_k^\delta) - f_k - \beta (K^* g - K^* K f_k) \\ &= \underbrace{f_k^\delta - f_k}_{=d_k^\delta} + \beta K^* (g^\delta - g) - \beta K^* K \underbrace{(f_k^\delta - f_k)}_{=d_k^\delta} \\ &= (I - \beta K^* K) d_k^\delta + \beta K^* (g^\delta - g), \quad d_0^\delta = 0. \end{aligned} \quad (4.8)$$

Kolmioepäyhtälön ja induktion avulla saadaan

$$\begin{aligned} \|d_{k+1}^\delta\| &\leq \|(I - \beta K^* K) d_k^\delta\| + \|\beta K^* (g^\delta - g)\| \\ &\leq \underbrace{\|I - \beta K^* K\|}_{<1} \|d_k^\delta\| + \beta \|K\| \underbrace{\|g^\delta - g\|}_{=\delta} \\ &\leq \|d_k^\delta\| + \beta \|K\| \delta \\ &\leq k\beta \|K\| \delta, \end{aligned} \quad (4.9)$$

sillä kun $d_0^\delta = 0$, niin

$$\begin{aligned} \|d_1^\delta\| &\leq \beta \|K\| \delta, \\ \|d_2^\delta\| &\leq \underbrace{\|d_1^\delta\|}_{\leq \beta \|K\| \delta} + \beta \|K\| \delta \leq 2\beta \|K\| \delta, \\ &\vdots \end{aligned}$$

Edelleen

$$\|f_k^\delta - K^\dagger g\| = \|f_k^\delta - f_k + f_k - K^\dagger g\|$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 93 / 108

Takaisin

Full Screen

Sulje

Poistu



Sisältö:

1. Johdanto
2. Yleistetyt käänteisluvut
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu

◀ ▶

◀ ▶

Sivu 94 / 108

Takaisin

Full Screen

Sulje

Poistu

$$\begin{aligned}
&\leq \|f_k - K^\dagger g\| + \|f_k^\delta - f_k\| \\
&\leq \|f_k - K^\dagger g\| + \mathcal{O}(k\delta) .
\end{aligned} \tag{4.10}$$

Valitaan algoritmin *pysäytys-* eli *terminointisääntö* $k = k(\delta)$ siten, että kun $\delta \rightarrow 0$, niin $\|f_k^\delta - f_k\| \rightarrow 0$. Tämä toteutuu, kun $k\delta \rightarrow 0$, eli kun $\delta \rightarrow 0$, jolloin voidaan valita esimerkiksi $k(\delta) = c_0/\sqrt{\delta}$.

4.3. Konjugaattigradientti-iteraatiot

Olkoon $A \in \mathbb{R}^{N \times N}$ symmetrinen ja positiividefiniitti. Nyt yhtälöryhmän $Af = g$ ratkaisu minimoi myös neliöllisen funktionaalin (laskuharjoitukset)

$$\phi(f) = \frac{1}{2} \|f\|_A^2 - \langle f, g \rangle . \tag{4.11}$$

Siten yhtälöryhmän ratkaisua voidaan hakea numeerisen minimoinnin keinoin. Ideana on, että suunta voi olla “halvempi” laskea kuin yhtälöryhmän ratkaisu $f = A^{-1}g$.

Steepest descent

Yleinen ensimmäisen kertaluvun menetelmä on niin sanottu nopeimman laskun eli *steepest descent (SD)* menetelmä. Funktionaali (4.11) pienenee pisteessä f_k nopeimmin negatiivisen gradientin $-\nabla\phi(f_k) = g - Af_k$ suuntaan. Koska tämä suure on myös menetelmän residuaali, merkitään $r_k = g - Af_k$. SD-menetelmässä muodostetaan rekursio

$$f_k = f_{k-1} + \alpha_k r_{k-1} \tag{4.12}$$

ja siirrytään negatiivisen gradientin r_{k-1} suuntaan niin pitkälle, kunnes *yksiulotteisen haun* minimi saavutetaan, eli määrätään $\alpha_k \in \mathbb{R}$ siten, että $\phi(f_{k-1} + \alpha_k r_{k-1})$ minimoituu. Tämä minimi saavutetaan, kun

$$\alpha_k = \|r_{k-1}\|^2 \|r_{k-1}\|_A^{-2} . \tag{4.13}$$

Voidaan osoittaa, että SD-menetelmän globaali konvergenssi on hidas, jos $\kappa(A) = \lambda_{\max}/\lambda_{\min}$ on suuri, mikä on käänteisongelmien yhteydessä yleensä tilanne.



A-konjugaattigradientsuunnat

Tavoitteena on valita hakusuunnat jollain toisella tavalla, eli korvata suunnat r_k joillakin muilla suunnilla p_k siten, että konvergenssi olisi SD-menetelmää nopeampi. Nyt

$$\begin{aligned}
\phi(f_{k-1} + \alpha_k p_k) &= \frac{1}{2} \langle f_{k-1} + \alpha_k p_k, A(f_{k-1} + \alpha_k p_k) \rangle - \langle f_{k-1} + \alpha_k p_k, g \rangle \\
&= \frac{1}{2} [\langle f_{k-1}, Af_{k-1} \rangle + \alpha_k \langle f_{k-1}, Ap_k \rangle + \alpha_k \langle p_k, Af_{k-1} \rangle + \alpha_k^2 \langle p_k, Ap_k \rangle] \\
&\quad - \langle f_{k-1}, g \rangle - \alpha_k \langle p_k, g \rangle \quad \text{ja} \\
\frac{d\phi}{d\alpha_k} &= \langle Af_{k-1}, p_k \rangle + \alpha_k \langle p_k, Ap_k \rangle - \langle p_k, g \rangle \\
&= \underbrace{\langle Af_{k-1} - g, p_k \rangle}_{=-r_{k-1}} + \alpha_k \|p_k\|_A^2 = 0 \\
\Leftrightarrow \alpha_k &= \langle r_{k-1}, p_k \rangle \|p_k\|_A^{-2},
\end{aligned}$$

joten yksiulotteisen minimoinnin tuloksena saadaan

$$\alpha_k = \operatorname{sol} \min_{\alpha_k} \phi(f_{k-1} + \alpha_k p_k) \quad (4.14)$$

$$= \langle p_k, r_{k-1} \rangle \|p_k\|_A^{-2}, \quad (4.15)$$

jolloin

$$\phi(f_k) = \phi(f_{k-1}) - \frac{1}{2} \langle p_k, r_{k-1} \rangle \|p_k\|_A^{-2}. \quad (4.16)$$

Ongelmana on, kuinka suunnat p_k tulisi valita, jotta algoritmi suppenisi mahdollisimman nopeasti. Jos esimerkiksi $p_k \perp r_{k-1}$, ei modifioitu algoritmi suppene lainkaan.

Oletetaan, että kyetään valitsemaan lineaarisesti riippumattomat (LI) suunnat p_ℓ , $\ell = 1, \dots, k$, siten, että

$$f_k = \operatorname{sol} \min_{f_k \in \mathcal{P}_k} \phi(f_k), \quad (4.17)$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 95 / 108

Takaisin

Full Screen

Sulje

Poistu



missä $\mathcal{P}_k = \text{span}\{p_1, \dots, p_k\}$. Koska $\{p_\ell\}$ on LI, on $\mathcal{P}_N = \mathbb{R}^N$ ja $Af_N = g$, joten tällainen algoritmi suppenisi oikeaan ratkaisuun N :llä askeleella. Pyritään määräämään suureet α_k ja p_k siten, että

$$f_k = \text{sol} \min_{\substack{\alpha_k \\ f_k \in \mathcal{P}_k}} \phi(f_{k-1} + \alpha_k p_k). \quad (4.18)$$

Olkoon nyt $P_k = (p_1, \dots, p_k) \in \mathbb{R}^{N \times k}$ valittu ja $\{p_k\}$ LI. Jos $f_{k-1} \in \mathcal{R}(P_{k-1})$, on $f_k = P_{k-1}y_k + \alpha_k p_k$, missä vektorin $y_k \in \mathbb{R}^{k-1}$ elementit ovat f_{k-1} :n koordinaatit kannan $\{p_1, \dots, p_{k-1}\}$ suhteen. Silloin

$$\phi(f_k) = \phi(P_{k-1}y_k) + \alpha_k \langle P_{k-1}y_k, Ap_k \rangle + \frac{\alpha_k^2}{2} \|p_k\|_A - \alpha_k \langle p_k, g \rangle. \quad (4.19)$$

Pyritään eliminoimaan oikean käden toinen termi, eli valitaan p_k siten, että $\langle P_{k-1}y_k, Ap_k \rangle = \langle y_k, P_{k-1}^* Ap_k \rangle = 0$ kaikille y_k , mikä toteutuu, kun

$$p_k \in \mathcal{N}(P_{k-1}^* A). \quad (4.20)$$

Tästä seuraa, että $p_k \in \mathcal{R}(A^* P_{k-1})^\perp = \mathcal{R}(AP_{k-1})^\perp = \text{span}\{Ap_1, \dots, Ap_{k-1}\}^\perp$. Ehtoa (4.20) kutsutaan *konjugaattigradienttiehdoksi*. Tällöin

$$\min_{\alpha_k, \alpha_k} \phi(f_k) = \min_{y_k} \phi(P_{k-1}y_k) + \min_{\alpha_k} \left\{ \frac{\alpha_k^2}{2} \|p_k\|_A^2 - \alpha_k \langle p_k, g \rangle \right\}, \quad (4.21)$$

jossa molemmat minimoinnit voidaan suorittaa toisistaan riippumatta. Koska $f_{k-1} \in \mathcal{R}(P_{k-1})$, on konjugaattigradienttiehdon toteutuessa $\langle p_k, Af_{k-1} \rangle = 0$, jolloin

$$\alpha_k = \langle p_k, g \rangle \|p_k\|_A^{-2} \quad (4.22)$$

$$= \langle p_k, r_{k-1} \rangle \|p_k\|_A^{-2}. \quad (4.23)$$

Voidaan osoittaa, että jos $r_{k-1} \neq 0$, niin on olemassa $p_k \in \mathcal{N}(P_{k-1}^* A)$, joka ei kuitenkaan ole yksikäsitteinen. Valitaan näistä p_k se, jolle $\|p_k - r_{k-1}\|$ on pienin. Tällöin

$$r_k = r_{k-1} - \alpha_k Ap_k, \quad (4.24)$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 96 / 108

Takaisin

Full Screen

Sulje

Poistu



$$\langle p_k, r_k \rangle = 0, \quad (4.25)$$

$$\mathcal{P}_k = \text{span}\{r_0, \dots, r_{k-1}\} = \text{span}\{g, Ag, \dots, A^{k-1}g\} \quad \text{ja} \quad (4.26)$$

$$\langle r_k, r_\ell \rangle = 0, \quad k \neq \ell. \quad (4.27)$$

Siten voidaan kirjoittaa $p_k = r_{k-1} + \beta_k p_{k-1}$ ja edelleen osoittaa, että

$$\beta_k = -\langle p_{k-1}, Ar_{k-1} \rangle \|p_{k-1}\|_A^{-2} \quad \text{ja} \quad (4.28)$$

$$\alpha_k = \|r_{k-1}\|_2^2 \|p_k\|_A^{-2}. \quad (4.29)$$

Kun käytetään hyväksi algoritmissa aiemmin laskettuja välituloksia, saadaan CG-algoritmiksi

$k = 0, f_0 = f_0, r_0 = g - Af_0$

```

while  $r_k \neq 0$ 
   $k = k + 1$ 
  if  $k = 1$ 
     $p_1 = r_0$ 
  else
     $\beta_k = r_{k-1}^T r_{k-1} / r_{k-2}^T r_{k-2}$ 
     $p_k = r_{k-1} + \beta_k p_{k-1}$ 
  end
   $q_k = Ap_k$ 
   $\alpha_k = r_{k-1}^T r_{k-1} / p_k^T q_k$ 
   $f_k = f_{k-1} + \alpha_k p_k$ 
   $r_k = r_{k-1} - \alpha_k q_k$ 
end

```

end
 $f = f_k$

Kuten algoritmista nähdään, raskain laskutoimitus, eli matriisi-vektori -tulo, lasketaan vain kerran iteraatiokierroksessa.

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 97 / 108

Takaisin

Full Screen

Sulje

Poistu



4.4. Kaczmarz-iteraatiot

Kaczmarz-iteraatio yhtälöryhmän $g = Kf$ ratkaisemiseksi perustuu tämän tarkasteluun riveittäin seuraavasti. Olkoon K_ℓ matriisin K ℓ :s rivi. Jokainen yhtälöryhmän yhden rivin määräämä yhtälö on luonnollisesti alideterministinen ja ratkaisuja f , jotka toteuttavat yhtälön

$$g_\ell = \langle K_\ell, f \rangle \quad (4.30)$$

on äärettömän monta. Ratkaisut, jotka toteuttavat yhtälön (4.30), muodostavat moniston. Olkoon P_ℓ projektori monistolle (4.30). Kaczmarz-iteraatiossa muodostetaan ratkaisujono $f^{(\ell)} = P_\ell f^{(\ell-1)}$.

4.4.0.1. PARALLEL BEAM -TOMOGRAFIA

Tarkastellaan aluksi, kuinka monistot liittyvät parallel beam -tomografiaongelman muodostamiseen. Menetelmää, jossa monistoja “ei säilötä”, vaan jokaisen havainnon yhteydessä iteroidaan yksi kierros Kaczmarz-iteraatiota, kutsutaan ART-algoritmiksi (*Algebraic Reconstruction Technique*).

Olkoon $f(x) > 0$, $x \in \Omega \subset \mathbb{R}^2$ ja Ω kompakti. Tarkastellaan kuvan 4.1 osoittamaa tilannetta. Olkoon g viivaintegraali $g = \int_\gamma f(x)$, joka voidaan tulkita formaalisti (normeeraamattomaksi) projektioksi $g = \langle p_\gamma, f \rangle$, missä p_γ on distribuutio. Diskretoidaan ongelma ja tarkastellaan M :ää projektiota. Jaetaan Ω pistevieraisiin osajoukkoihin (tyypillisesti taso jaetaan neliöihin) Ω_ℓ , jotka peittävät Ω :n, eli

$$\begin{aligned} \bigcup_{\ell=1}^M \Omega_\ell &= \Omega, \\ \int_{\Omega} \chi_\ell(x) \chi_k(x) &= 0, \quad \ell \neq k \quad \text{ja} \\ \chi_\ell(x) &= \begin{cases} 1, & x \in \Omega_\ell \\ 0, & \text{muulloin,} \end{cases} \end{aligned}$$

missä $\chi_\ell(x)$ on joukon Ω_ℓ indikaattorifunktio. Approksimoidaan nyt funktiota $f(x)$ funktiolla,

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 98 / 108

Takaisin

Full Screen

Sulje

Poistu



Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



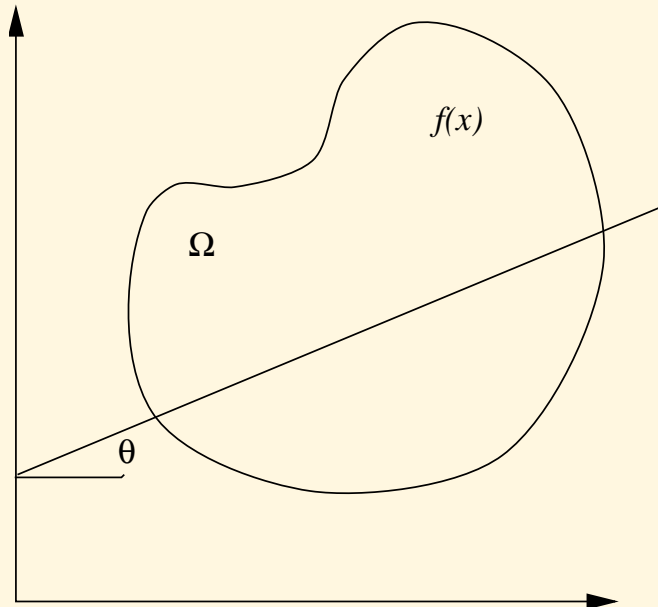
Sivu 99 / 108

Takaisin

Full Screen

Sulje

Poistu



Kuva 4.1: Projektiot hypertasoille.



joka on paloittain vakio

$$f(x) \approx \sum_{j=1}^M f_j \chi_j(x) .$$

Oletetaan, että on havaittu projektiot (viivaintegraalit) $g_\ell = \int_{\gamma_\ell} f(x)$, missä γ_ℓ on projektiota vastaava suora. Nyt voidaan kirjoittaa approksimaatio

$$g_\ell \approx \langle K_\ell, f \rangle \quad \text{ja} \quad (4.31)$$

$$K_\ell(j) = \int_{\gamma_\ell} \chi_j(x) , \quad (4.32)$$

missä $f, K_\ell \in \mathbb{R}^N$. Alkio $K_\ell(j)$ voidaan siis tulkita suoran γ_ℓ elementtiä Ω_j leikkaavan suoran pituudeksi.

Edellä esitetty diskretoitu ongelma voidaan kirjoittaa yhtälöryhmän muotoon

$$g = Kf \quad \text{ja} \quad (4.33)$$

$$K = [K_1, \dots, K_M]^T , \quad (4.34)$$

joka voi olla alideterministinen, eli $M < N$, sekä lisäksi useissa sovelluksissa hyvin suuri, mutta usein myös hyvin harva matriisi.

4.4.1. Projektiot monistoille

Tarkastellaan monistoa $\mathcal{V}_* = \{f \mid \langle v, f \rangle = g\}$, $v \neq 0$, ja merkitään $\mathcal{V} = \text{span}\{v\}$, kuva 4.2. Vektorit $f \in \mathcal{V}_*$ voidaan kirjoittaa muotoon $f = f_1 + f_2$, $f_1 \in \mathcal{V}$ ja $f_2 \in \mathcal{V}^\perp$. Nyt $f_1 = cv$ ja

$$\langle v, f_1 + f_2 \rangle = \langle v, f_1 \rangle = c \langle v, v \rangle = g , \quad (4.35)$$

joten $c = g \|v\|^{-2}$. Siten monisto voidaan esittää muodossa $\mathcal{V}_* = \{g \|v\|^{-2} v + f_2 \mid f_2 \in \mathcal{V}^\perp\}$. Olkoon $x \in \mathcal{H}$ mielivaltainen, jolloin on olemassa yksikäsitteinen $P_{\mathcal{V}_*} x \in \mathcal{V}_*$ siten, että

$$\|x - P_{\mathcal{V}_*} x\| = \min_{y \in \mathcal{V}_*} \|x - y\| . \quad (4.36)$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 100 / 108

Takaisin

Full Screen

Sulje

Poistu



Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



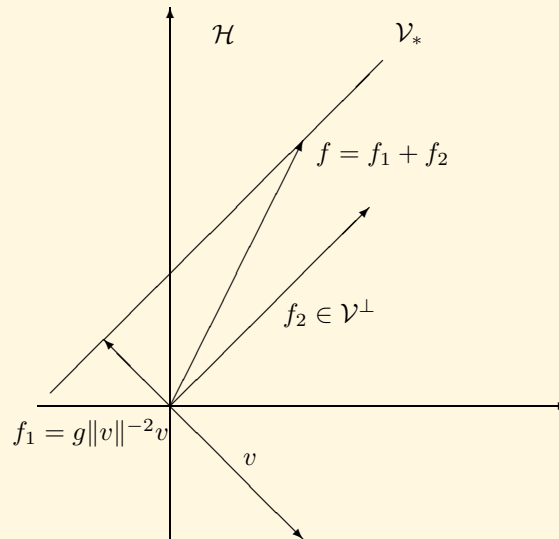
Sivu 101 / 108

Takaisin

Full Screen

Sulje

Poistu



Kuva 4.2: Monisto $\mathcal{V}_* = \{f \mid \langle v, f \rangle = g\}$.



Moniston \mathcal{V}_* vektorit ovat muotoa $w_1 + w_0$, missä $w_1 \in \mathcal{V}_*$ ja $w_0 \in \mathcal{V}^\perp$. Siten funktiolla

$$h(t) = \|x - (P_{\mathcal{V}_*}x + tw)\|^2 \quad (4.37)$$

on oltava minimi pisteessä $t = 0$ kaikille kiinteille $w \in \mathcal{V}^\perp$. Asetetaan $h'(t) = 0$, jolloin saadaan

$$h'(0) = \frac{d}{dt} \{ \|x - (P_{\mathcal{V}_*}x + tw)\|^2 \} \Big|_{t=0} \quad (4.38)$$

$$= \frac{d}{dt} \{ \langle x - P_{\mathcal{V}_*}x - tw, x - P_{\mathcal{V}_*}x - tw \rangle \} \Big|_{t=0} \quad (4.39)$$

$$= 2 \langle w, x - (P_{\mathcal{V}_*}x - tw) \rangle \Big|_{t=0} \quad (4.40)$$

$$= 2 \langle w, x - P_{\mathcal{V}_*}x \rangle = 0 \quad (4.41)$$

kaikille w , kun $t = 0$. Siten $x - P_{\mathcal{V}_*}x \in \mathcal{V}^{\perp\perp}$, joten $x - P_{\mathcal{V}_*}x = \beta v$, kun $\beta \in \mathbb{R}$, kuva 4.3. Nyt $P_{\mathcal{V}_*}x = x - \beta v \in \mathcal{V}_*$ ja

$$g = \langle v, P_{\mathcal{V}_*}x \rangle = \langle v, x \rangle - \beta \|v\|^2, \quad (4.42)$$

joten

$$\beta = (\langle v, x \rangle - g) \|v\|^{-2} \quad \text{ja} \quad (4.43)$$

$$P_{\mathcal{V}_*}x = x - (\langle v, x \rangle - g) \|v\|^{-2} v. \quad (4.44)$$

Voidaan osoittaa, että $\|P_{\mathcal{V}_*}x_1 - P_{\mathcal{V}_*}x_2\| \leq \|x_1 - x_2\|$, joten projektori $P_{\mathcal{V}_*}$ on ei-laaajentava (vrt. kutistava). Yhtälöryhmän $Kf = g$ ratkaisuille on siis oltava voimassa

$$\langle K_\ell, f \rangle = g_\ell, \quad \ell = 1, \dots, M, \quad (4.45)$$

eli f kuuluu jokaiseen parien (K_ℓ, g_ℓ) määräämään monistoon $\mathcal{V}_{*,\ell}$ eli

$$f \in \bigcap_{\ell=1}^M \mathcal{V}_{*,\ell}. \quad (4.46)$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 102 / 108

Takaisin

Full Screen

Sulje

Poistu



Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



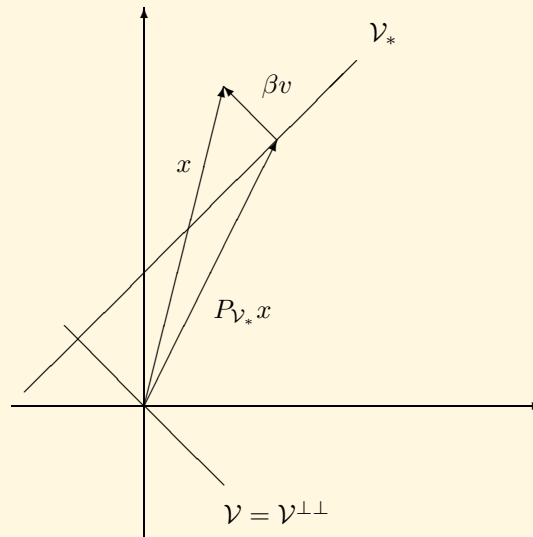
Sivu 103 / 108

Takaisin

Full Screen

Sulje

Poistu

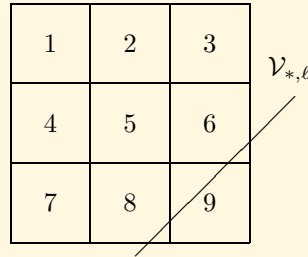


Kuva 4.3: Vektorin x projektiio $P_{\mathcal{V}_*}x$ monistolle \mathcal{V}_* .



Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä



$$K_\ell = [0 \ 0 \ 0 \ 0 \ 0 \ 1/3 \ 0 \ 1/3 \ 1]$$

$$f \leftarrow f + \underbrace{(g_\ell - K_\ell^T f) \|K_\ell\|^{-2} K_\ell}_{\in \mathbb{R}}$$

Kuva 4.4: Ratkaisualkioiden päivittyminen Kaczmarz-iteraatioissa.

Kaczmarz-iteraatioiden ideana on valita alkuarvo, projisoida tämä rekursiivisesti monistoille $\mathcal{V}_{*,\ell}$ (ja jatkaa alusta) siten, että

$$f^{(\ell+1)} = P_\ell f^{(\ell)}, \tag{4.47}$$

missä P_ℓ on projektori monistolle $\mathcal{V}_{*,\ell}$. Algoritmi saa nyt muodon

```

k = 0, f = f_0
while (lopetusehto)
  for l = 1, ..., M
    f = P_l f = f + (g_l - K_l^T f) \|K_l\|^{-2} K_l
  end
end

```

Selvästi nähdään, että projektiovektorit päivittävät vain niitä f :n alkioita, jotka ovat projektiossa “mukana”, kuva 4.4.

Etusivu



Sivu 104 / 108

Takaisin

Full Screen

Sulje

Poistu



4.5. Esimerkki iteratiivisista menetelmistä

Tarkastellaan dekonvoluutio-ongelmaa, jota kuvaa integraaliyhtälö

$$g(x) = \int_0^1 e^{-\xi|x-y|} f(y) dy$$

missä $\xi > 0$. Olkoon (kohinaton) data

$$g(x) = 2\xi^{-1}x - \xi^{-1}e^{-\xi(1-x)} + \xi^{-2}(e^{-\xi x} - e^{-\xi(1-x)}),$$

jolloin integraaliyhtälön yksikäsitteinen ratkaisu on

$$f(y) = y$$

ja ratkaisun ongelmallisuus riippuu parametrasta ξ . Data g diskretoidaan tasaväliseen hilaan $x_1 = 0, x_2 = h, \dots, x_N = 1$, missä $h = 0.01$ ja siten $N = h^{-1} + 1 = 101$. f diskretoidaan hilaan $y_1 = 0, y_2 = \Delta y, \dots, y_M = 1$, missä Δy saa eri arvoja ja $M = (\Delta y)^{-1} + 1$. Diskretoidaan K seuraavasti

$$K = \begin{pmatrix} w_1 \exp(-\xi |x_1 - y_1|) & \dots & w_M \exp(-\xi |x_1 - y_M|) \\ w_1 \exp(-\xi |x_2 - y_1|) & \dots & w_M \exp(-\xi |x_2 - y_M|) \\ \vdots & & \vdots \\ w_1 \exp(-\xi |x_N - y_1|) & \dots & w_M \exp(-\xi |x_N - y_M|) \end{pmatrix},$$

jossa

$$w_k = \begin{cases} \Delta y/2 & k = 1, N \\ \Delta y & k = 2, \dots, N-1 \end{cases}.$$

Ratkaisuhilan dimensio on $M = 81$ ja datavektorin pituus on $N = 101$. Todellinen ratkaisu on $f(y) = y$ ja additiivisen kohinan varianssi oli

$$\|e\| \approx 10^{-3} \|g\|.$$

Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



Sivu 105 / 108

Takaisin

Full Screen

Sulje

Poistu



Sisältö:

1. Johdanto
2. Yleistetyt käänteiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



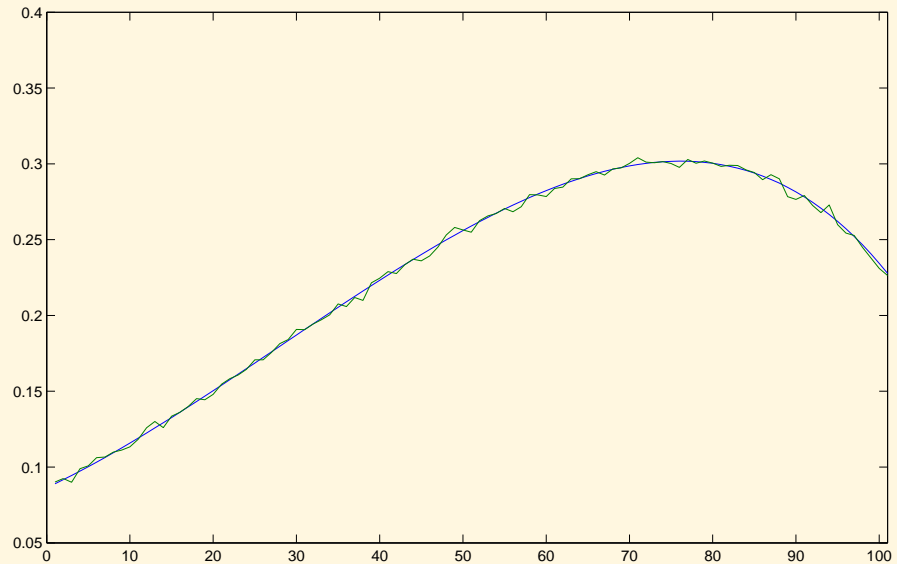
Sivu 106 / 108

Takaisin

Full Screen

Sulje

Poistu



Kuva 4.5: Kohinaton ja kohinainen data. [\[KOODI\]](#)



Ajetaan kutakin menetelmää sata kierrosta (Kaczmarz-iteraatioita sata kertaa kaikki rivit yksitellen) parametrin ξ arvoilla $\xi = 3, 10$ ja 50 . Kohinaton ja kohinainen data tapauksessa $\xi = 3$ on esitetty kuvassa 4.5.

Landweber-iteraatioissa kertoimena käytettiin $0.95\beta_{\max}$. Estimointivirheiden normit $\|f_k - f\|$ sekä näiden mielessä parhaat estimaatit on esitetty kuvassa 4.6. ART-algoritmin virheistä on tulostettu joka kymmenes.

Huomionarvoista on, että CG ja Landweber iteraatioiden antamat parhaat estimaatit ovat lähes yhteneviä, mutta CG “saavuttaa” parhaan estimaatin pienemmällä määrällä kierroksia kuin Landweber. Myös optimaalisen iteraatiokierrosten määrän riippuvuus operaattorin K huonokuntoisuudesta (tässä tapauksessa parametrissa ξ) on selvästi havaittavissa.

Tehtäviä

Sisältö:

1. *Johdanto*
2. *Yleistetyt käänteiskuvaukset*
3. *Tihonov-regularisointi*
4. *Iteratiiviset menetelmät*
5. *Tehtäviä*

Etusivu



Sivu 107 / 108

Takaisin

Full Screen

Sulje

Poistu



Sisältö:

1. Johdanto
2. Yleistetyt käänneiskuvaukset
3. Tihonov-regularisointi
4. Iteratiiviset menetelmät
5. Tehtäviä

Etusivu



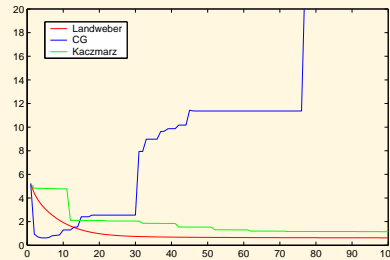
Sivu 108 / 108

Takaisin

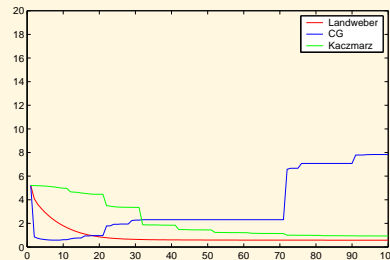
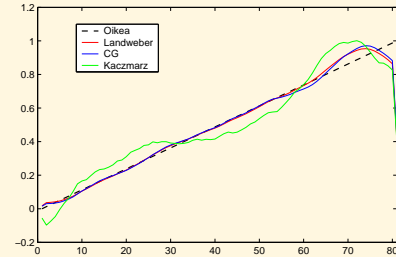
Full Screen

Sulje

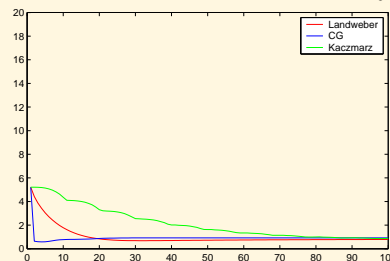
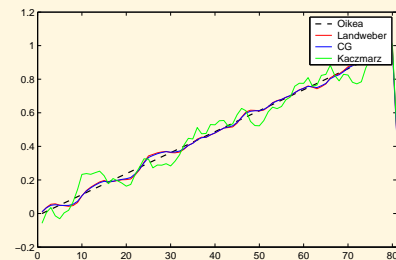
Poistu



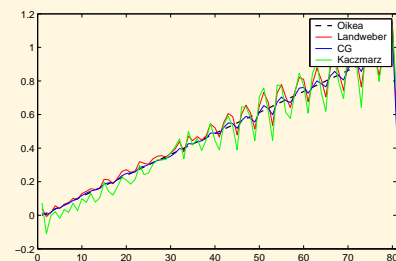
$\xi = 3$



$\xi = 10$



$\xi = 50$



Kuva 4.6: Estimointivirheiden normit $\|f_k - f\|$ kierroksen funktiona (vasen sarake) sekä näiden mielessä parhaat estimaatit (oikea sarake) sekä todellinen f (katkoviiva). Paksut viivat: Kaczmarz (ART), keskipaksut: konjugaattigradiendi ja ohuet viivat: Landweber. [\[KOODI\]](#)



Sisältö:

1. *Johdanto*
2. *Yleistetyt
käänteiskuvaukset*
3. *Tihonov-
regularisointi*
4. *Iteratiiviset
menetelmät*
5. *Tehtäviä*

Etusivu



Sivu 109 / 108

Takaisin

Full Screen

Sulje

Poistu

Kappaleen 1 tehtäviä
Kappaleen 2 tehtäviä
Kappaleen 3 tehtäviä
Kappaleen 4 tehtäviä