Instytut Badań Systemowych
POLSKIEJ AKADEMII NAUK

Piotr Suchomski

SYNTEZA ALGORYTMÓW ODPORNEGO STEROWANIA W CZASIE DYSKRETNYM

Piotr Suchomski

SYNTEZA ALGORYTMÓW ODPORNEGO STEROWANIA W CZASIE DYSKRETNYM

INSTYTUT BADAŃ SYSTEMOWYCH • POLSKIEJ AKADEMII NAUK

Seria: BADANIA SYSTEMOWE

tom 38

Redaktor naukowy:

Prof. dr hab. inż. Jakub Gutenbaum

Warszawa 2004

Piotr Suchomski

SYNTEZA ALGORYTMÓW ODPORNEGO STEROWANIA W CZASIE DYSKRETNYM

Publikację opiniowali do druku:

Prof. dr hab. inż. Mikołaj Busłowicz

Doc. dr hab. inż. Piotr Kulczyki (prof. PK)

Wydawca: Instytut Badań Systemowych PAN

ul. Newelska 6 01-447 Warszawa

Sekcja Informacji Naukowej i Wydawnictw

tel. 837-68-22

email: biblioteka@ibspan.waw.pl

ISBN 83-85847-94-4

ISSN 0208-8029

Dodatek A

Uwarunkowanie liniowych zadań

Uwarunkowanie liniowego zadania

Rozważmy równanie $(A+\epsilon\tilde{A})x(\epsilon) = b+\epsilon\tilde{b}$, $x(0) = x \neq 0_n$, gdzie A, $\tilde{A} \in \mathbb{R}^{n\times n}$ oraz b, $\tilde{b} \in \mathbb{R}^n$, przy czym $b \neq 0_n$. W przypadku nieosobliwej macierzy A odwzorowanie $x : \mathbb{R} \to \mathbb{R}^n$ jest różniczkowalne w otoczeniu $\epsilon = 0$, co więcej: $\dot{x}(0) = A^{-1}(\tilde{b} - \tilde{A}x)$. Analizując szerg Taylora tego odwzorowania, otrzymujemy oszacowanie $\|x(\epsilon) - x\|/\|x\| \leq \kappa(A)(\rho_A + \rho_b) + \mathcal{O}(\epsilon^2)$, gdzie $\kappa(A) = \|A\|\,\|A^{-1}\|$ jest pewnym wskaźnikim uwarunkowania macierzy A, zaś wielkości $\rho_A = |\epsilon|\|\tilde{A}\|/\|A\|$ oraz $\rho_b = |\epsilon|\|\tilde{b}\|/\|b\|$ reprezentują względne charakterystyki odpowiednich zaburzeń (Björck [32], Golub i Van Loan [139], Higham [177], Kato [214], Rump [345, 346], Stewart [381]). Przyjmuje się przy tym, że macierzowa norma stosowana w przestrzeni $\mathbb{R}^{n\times n}$ jest zgodna z normą w przestrzeni $\mathbb{R}^n$. Rozwiązywanie nieosobliwego liniowego problemu należy zatem do klasy zadań ciągłych lokalnie w sensie Lipschitza. Najchętniej używanymi tak zdefiniowanymi wskaźnikami uwarunkowania są: $\kappa_2(A) = \|A\|_2\,\|A^{-1}\|_2 = \sigma_{\max}(A)/\sigma_{\min}(A)$ odpowiadający macierzowej spektralnej normie indukowanej przez normę $\|\cdot\|_2$ w $\mathbb{R}^n$, a także $\kappa_\infty(A) = \|A\|_\infty\,\|A^{-1}\|_\infty$ przyporządkowany macierzowej normie $\|A\|_\infty = \max_i \sum_j |\bar{a}_{ij}|$ indukowanej przez wektorową normę $\|\cdot\|_\infty$. Wyznaczanie wskaźników uwarunkowania jest w ogólnym przypadku złożonym zadaniem obliczeniowym (Dhillon [88], Higham [177], Stewart [381]). Duże znaczenie mają zatem takie algorytmy oceny wskaźnika uwarunkowania, w których nie jest wymagane odwracanie macierzy. Można tu przykładowo wymienić procedury xyyCON z pakietu LAPACK (Anderson *et al.* [4]), stosowane przy szacowaniu $\kappa_1(A)$. Dla dowolnej macierzowej normy zachodzi

$$\kappa(A) = \lim_{\epsilon \to 0} \sup_{\|\tilde{A}\| \le \epsilon \|A\|} \frac{\|(A+\tilde{A})^{-1} - A^{-1}\|}{\epsilon} \frac{1}{\|A^{-1}\|}$$

co oznacza, że wskaźnik uwarunkowania $\kappa(A)$ jest pewną unormowaną charakterystyką pochodnej Frécheta odwzorowania, które danej macierzy przyporządkowuje jej odwrotność (Rice [335]). Warto podkreślić, że chociaż wszystkie macierzowe normy w $\mathbb{R}^{n\times n}$ są równoważne, to jednak pytanie o wybór normy 'właściwej' dla danego szczegółowego problemu jest pytaniem ważnym (Golub i Van Loan [139], Grcar [150], Stewart [381]).

Lemat D.1 (o ocenie względnego błędu rozwiązania nieosobliwego zadania liniowego z niestrukturalizowalnymi zaburzeniami; Higham [177]). *Niech $Ax = b$, $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$ oraz $(A+\tilde{A})y = b + \tilde{b}$, $\tilde{A} \in \mathbb{R}^{n\times n}$, $\tilde{b} \in \mathbb{R}^n$, przy czym A jest macierzą nieosobliwą, $\|\tilde{A}\| \le \epsilon\|E\|$ oraz $\|\tilde{b}\| \le \epsilon\|f\|$ dla pewnych $E \in \mathbb{R}^{n\times n}$, $f \in \mathbb{R}^n$. Jeżeli $\epsilon\|A^{-1}\|\|E\| < 1$, wtedy dla dowolnej wektorowej normy oraz odpowiedniej indukowanej normy macierzowej*

$$\frac{\|x-y\|}{\|x\|} \le \frac{\epsilon}{1-\epsilon\|A^{-1}\|\|E\|} \kappa_{E,f}(A,x)$$

gdzie wskaźnik uwarunkowania $\kappa_{E,f}(A,x)$ zdefiniowany jako

$$\kappa_{E,f}(A,x) = \lim_{\epsilon \to 0} \sup \left\{ \frac{\|\tilde{x}\|}{\epsilon\|x\|} : (A+\tilde{A})(x+\tilde{x}) = b+\tilde{b},\ \|\tilde{A}\| \le \epsilon\|E\|,\ \|\tilde{b}\| \le \epsilon\|f\| \right\}$$

dany jest wzorem

$$\kappa_{E,f}(A,x) = \frac{\|A^{-1}\|\|f\|}{\|x\|} + \|A^{-1}\|\|E\|\,. \ \square$$

Lemat D.2 (o ocenie względnego błędu rozwiązania nieosobliwego zadania liniowego ze strukturalizowalnymi zaburzeniami; Higham [177]-[178], Rump [345]). *Niech $Ax = b$, $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$ oraz $(A+\tilde{A})y = b+\tilde{b}$, $\tilde{A} \in \mathbb{R}^{n\times n}$, $\tilde{b} \in \mathbb{R}^n$, przy czym A jest macierzą nieosobliwą, $|\tilde{A}| \le \epsilon E$ oraz $|\tilde{b}| \le \epsilon f$ dla pewnych $E \ge 0_{n\times n} \in \mathbb{R}^{n\times n}$ oraz $f \ge 0_n \in \mathbb{R}^n$. Jeżeli $\epsilon\|\,|A^{-1}|E\,\|_\infty < 1$, wtedy*

$$\frac{\|x-y\|_\infty}{\|x\|_\infty} \le \frac{\epsilon}{1-\epsilon\|\,|A^{-1}|E\,\|_\infty} \kappa^\infty_{E,f}(A,x)$$

gdzie odpowiedni wskaźnik uwarunkowania $\kappa^\infty_{E,f}(A,x)$ zdefiniowany jako

$$\kappa^\infty_{E,f}(A,x) = \lim_{\epsilon\to 0} \sup \left\{ \frac{\|\tilde{x}\|_\infty}{\epsilon\|x\|_\infty} : (A+\tilde{A})(x+\tilde{x}) = b+\tilde{b},\ |\tilde{A}| \le \epsilon E,\ |\tilde{b}| \le \epsilon f \right\}$$

dany jest wzorem

$$\kappa_{E,f}^{\infty}(A,x) = \frac{\|\,|A^{-1}|(E|x|+f)\,\|_\infty}{\|x\|_\infty}.$$

Zamiast $\|\cdot\|_\infty$ *można stosować dowolną absolutną normę wektorową oraz odpowiednią indukowaną absolutną normę macierzową.* □

Symbole $|\cdot|$, $\leq$, $\geq$ oraz $>$ dotyczą teraz poszczególnych elementów macierzy oraz współrzędnych wektorów. Przyjmując w *lemacie D.2* macierz $E = |A|$ oraz wektor $f = |b|$, otrzymujemy wskaźniki uwarunkowania Skeela (Skeel [362]), nazywane także wskaźnikami Bauera-Skeela (Bauer [23], Björck [32], Peña [321], Stewart [381]):

$$\kappa_S(A,x) = \frac{\|\,|A^{-1}||A||b|\,\|_\infty}{\|x\|_\infty}, \qquad \kappa_S(A) = \|\,|A^{-1}||A|\,\|_\infty.$$

Ze względu na znaczenie wskaźnika $\kappa_\infty(A)$ istotne jest pytanie o to, jak dalece przez stosowne skalowanie danej macierzy A można poprawić uwarunkowanie odpowiedniego liniowego zadania (Chandrasekaran i Ipsen [55], Kiełbasiński i Schwetlick [219], Shapiro [358, 359], Van der Sluis [440, 441], Watson [460]).

Lemat D.3 (o wpływie skalowania na wskaźnik uwarunkowania $\kappa_\infty(A)$; Bauer [23, 24], Businger [49], Chan i Foulser [54], Higham [177], Stewart i Sun [383], Van der Sluis [440, 441]). *Niech* $A \in \mathbb{R}^{n\times n}$ *będzie nieosobliwą macierzą, zaś* $D_n^{>} = \{\mathrm{diag}\,\{d_i\}_{i=1}^n,\ d_i > 0, 1 \leq i \leq n\} \subset \mathbb{R}^{n\times n}$.

(*i*) *Zachodzi* $\inf_{D_c,D_r\in D_n^>} \kappa_\infty(D_rAD_c) \leq \rho(|A||A^{-1}|)$, *gdzie* $\rho(|A||A^{-1}|)$ *jest spektralnym promieniem macierzy* $|A||A^{-1}|$.

(*ii*) *Jeżeli* $|A|$ *oraz* $|A^{-1}|$ *są macierzami dodatnimi (Berman i Plemmons [29]), wtedy* $\min_{D_c,D_r\in D_n^>} \kappa_\infty(D_rAD_c) = \rho(|A||A^{-1}|)$.

(*iii*) *Jeżeli macierzy* $|A||A^{-1}|$ *oraz* $|A^{-1}||A|$ *nie można przez kongruencję* $P(\cdot)P^T$, *w której* $P \in \mathbb{R}^{n\times n}$ *jest macierzą permutacji, sprowadzić do macierzy o blokowej trójkątnej postaci, wtedy* $\min_{D_c,D_r\in D_n^>} \kappa_\infty(D_rAD_c)$ $= \rho(|A||A^{-1}|)$. *W takim przypadku minimum wskaźnika* $\kappa_\infty(D_rAD_c)$ *zapewniają odpowiednie skalujące macierze* $D_c = \mathrm{diag}\,(|A^{-1}|x_P)$ *oraz* $D_r = \mathrm{diag}\,(x_P)^{-1}$, *gdzie wektor* $x_P \in \mathbb{R}^n$ *o dodatnich współrzędnych,* $x_P > 0_n$, *jest prawym wektorem Perrona macierzy* $|A||A^{-1}|$, *czyli wektorem, dla którego* $|A||A^{-1}|x_P = \rho(|A||A^{-1}|)x_P$ *(Berman i Plemmons [29], Meyer [290]).* □

Numeryczna stabilność zadania liniowego

Uwarunkowania danego numerycznego zadania nie należy utożsamiać z oceną tak zwanej numerycznej stabilności algorytmów rozwiązywania tego zadania. W literaturze przedmiotu odnajdujemy różne definicje, zarówno formalne jak i nieformalne, pojęcia numerycznej stabilności (Björck [32], Bunch [46], Higham [177], Stewart [381], Van der Sluis, [442], Wilkinson [469, 470]). Rozważając uwarunkowanie liniowego zadania $Ax = b$, zdefiniujmy dwa wskaźniki (wsteczne błędy), charakteryzujące dane (niedokładne) rozwiązanie y tego zadania odpowiednio dla zaburzeń niestrukturalizowalnych (*normwise backward error*) oraz strukturalizowalnych (*component backward error*):

$$\begin{aligned} \eta_{E,f}(y) &= \min\{\epsilon : (A+\tilde{A})y = b+\tilde{b},\ \|\tilde{A}\| \le \epsilon\|E\|,\ \|\tilde{b}\| \le \epsilon\|f\|\} \\ \omega_{E,f}(y) &= \min\{\epsilon : (A+\tilde{A})y = b+\tilde{b},\ |\tilde{A}| \le \epsilon E,\ |\tilde{b}| \le \epsilon f\}, \\ & \qquad E \ge 0_{n\times n}, \quad f \ge 0_n\,. \end{aligned}$$

Zakładamy, że macierzowa norma użyta w tych definicjach jest normą indukowaną. Ponadto wprowadzamy pomocnicze oznaczenia $\eta^2_{E,f}(y)$ oraz $\eta^\infty_{E,f}(y)$, przyjmując w definicji wstecznego błędu $\eta_{E,f}(y)$ normę $\|\cdot\|_2$ oraz $\|\cdot\|_\infty$, odpowiednio.

Lemat D.4 (o wstecznych błędach rozwiązania nieosobliwego zadania liniowego; Oettli i Prager [308], Rigal i Gaches [336]). *Niech $r = b - Ay$ będzie resztowym wektorem odpowiadającym danemu oszacowaniu y rozwiązania nieosobliwego układu równań liniowych $Ax = b$, w którym $A \in \mathbb{R}^{n\times n}$ oraz $b \in \mathbb{R}^n$.*

(i) *W przypadku niestrukturalizowalnych zaburzeń $(A+\tilde{A})y = b+\tilde{b}$, gdzie $\|\tilde{A}\| \le \epsilon\|E\|$ oraz $\|\tilde{b}\| \le \epsilon\|f\|$, zachodzi*

$$\eta_{E,f}(y) = \frac{\|r\|}{\|E\|\|y\| + \|f\|}\,.$$

(ii) *W przypadku strukturalizowalnych zaburzeń $(A+\tilde{A})y = b+\tilde{b}$, gdzie $|\tilde{A}| \le \epsilon E$, $|\tilde{b}| \le \epsilon f\}$, przy czym $E \ge 0_{n\times n}$ oraz $f \ge 0_n$, mamy*

$$\omega_{E,f}(y) = \max_{1\le i\le n} \frac{|r_i|}{(E|y| + f)_i} \tag{A.1}$$

przy czym indeks i odnosi się do i-tej współrzędnej stosownego wektora, a wyrażenie $\xi/0$ interpretowane jest jako 0, gdy $\xi = 0$, zaś jako ∞ w każdym innym przypadku. □

Niech u_ε oznacza precyzję zmiennoprzecinkowej arytmetyki (Higham [177], Overton [310], Stewart [381]). Przykładowa MATLABowa implementacja standardu IEEE reprezentacji liczb o podwójnej precyzji zapewnia $u_\varepsilon \approx 2.22 \times 10^{-16}$. Przyjmijmy, że $y \in \mathbb{R}^n$ jest rozwiązaniem równania $Ax = b$, $A \in \mathbb{R}^{n\times n}$, uzyskanym przez zastosowanie danej numerycznej metody. Metodę tę uznajemy za numerycznie stabilną (w sensie danego wskaźnika), jeżeli odpowiednio:

$$\eta_{A,b}(y) = \mathcal{O}(u_\varepsilon), \qquad \omega_{|A|,|b|}(y) = \mathcal{O}(u_\varepsilon)$$

dla wszystkich 'dopuszczalnych' macierzy A oraz wektorów b. O danym rozwiązaniu y spełniającym powyższe kryteria mówimy, że jest rozwiązaniem numerycznie stabilnym. Zaletą takiego sformułowania definicji numerycznej stabilności jest to, że opiera się ono na wektorze resztowym $r = b - Ay$ możliwym do wyznaczenia jako pewna charakterystyka a posteriori rozwiązania y. Wymaganie numerycznej stabilności ze względu na wskaźnik $\omega_{|A|,|b|}(y)$ jest żądaniem silniejszym w zestawieniu z takim wymaganiem sformułowanym ze względu na $\eta_{A,b}(y)$. Stwierdziwszy zatem, że użyta metoda jest numerycznie stabilna w omawianym sensie, zyskujemy przekonanie, że y jest rozwiązaniem pewnego zaburzonego problemu dostatecznie 'bliskiego' problemowi postawionemu. W przypadku wstecznego błędu $\omega_{|A|,|b|}(y)$ można mówić o zaburzeniu wejściowych danych wywołanym ich konwersją na zmiennoprzecinkową postać. Dla metody numerycznie stabilnej ze względu na wsteczny błąd $\eta_{A,b}(y)$ obowiązuje górne ograniczenie względnego błędu $\|x - y\|/\|x\|$ wyznaczone przez $\kappa(A)u_\varepsilon$. Implikacja w drugą stronę nie musi jednak zawsze zachodzić. Możliwa jest bowiem taka sytuacja, w której obliczeniowa metoda dostarcza rozwiązania y opisanego ograniczoną wartością błędu $\|x-y\|/\|x\|$, lecz 'rząd' wstecznego błędu $\eta_{A,b}(y)$ istotnie różni się od 'rzędu' wielkości u_ε. Wynika stąd potrzeba 'autonomicznego' sformułowania definicji numerycznej stabilności także ze względu na błąd $\|x - y\|/\|x\|$ (podobnie jak poprzednio, stosowna definicja odnosząca się do strukturalizowalnych zaburzeń jest sformułowaniem silniejszym w stosunku do definicji opartej na charakterystykach zaburzeń niestrukturalizowalnych). Odpowiednie kryteria numerycznej stabilności mają teraz postać:

$$\frac{\|x-y\|}{\|x\|} = \mathcal{O}(\kappa(A)u_\varepsilon), \qquad \frac{\|x-y\|}{\|x\|} = \mathcal{O}(\kappa_S(A,x)u_\varepsilon)\,.$$

Relacje między podanymi kryteriami numerycznej stabilności ilustruje wzór (Higham [177], Wilkinson [469])

$$\begin{array}{ccc} \omega_{|A|,|b|}(y) = \mathcal{O}(u_\varepsilon) & \Rightarrow & \frac{\|x-y\|}{\|x\|} = \mathcal{O}(\kappa_S(A,x)u_\varepsilon) \\ \Downarrow & & \Downarrow \\ \eta_{A,b}(y) = \mathcal{O}(u_\varepsilon) & \Rightarrow & \frac{\|x-y\|}{\|x\|} = \mathcal{O}(\kappa(A)u_\varepsilon)\,. \end{array}$$

Sięgając po te kryteria numerycznej stabilności, w których korzystamy ze wskaźników uwarunkowania $\kappa(A)$ oraz $\kappa_S(A,x)$, musimy dokonać oszacowania względnego błędu danego rozwiązania y. W tym celu pomocne są *lematy D.1* oraz *D.2*, przy czym w miejsce ϵ kładziemy teraz wielkości wstecznych błędów $\eta_{A,b}(y)$ oraz odpowiednio $\omega_{|A|,|b|}(y)$, wyznaczone na podstawie znajomości resztowego wektora $r = b - Ay$. Ponieważ wektor x nie jest znany, wzór służący ocenie błędu rozwiązania y w oparciu o *lemat D.2* wymaga pewnej modyfikacji. Można w tym celu użyć standardowych oszacowań (Higham [177], Stewart [381]):

$$\begin{aligned} \frac{\|x-y\|_\infty}{\|y\|_\infty} &\le \omega_{|A|,|b|}(y)\kappa^\infty_{|A|,|b|}(A,y) \\ \frac{\|x-y\|_\infty}{\|y\|_\infty} &\le \frac{\|\,|A^{-1}||r|\,\|_\infty}{\|y\|_\infty}\,. \end{aligned}$$

Znane są efektywne przybliżone metody oceny górnego ograniczenia błędu $\|x-y\|_\infty/\|y\|_\infty$, w których nie występuje potrzeba wyznaczania odwrotności A^{-1} (mowa tu przykładowo o procedurach xyyRFS dostępnych w pakiecie LAPACK; Anderson *et al.* [4]). Przyjmujemy oznaczenia: $\varepsilon^p_{A,b}(y) = 2\eta^p_{A,b}(y)$ $\kappa_p(A)/(1-\eta^p_{A,b}(y)\kappa_p(A))$ dla $p = 2,\infty$, $\varepsilon^\omega_{A,b}(y) = \omega_{|A|,|b|}(y)\kappa^\infty_{|A|,|b|}(A,y)$ oraz $\varepsilon^{||}_{A,b}(y) = \|\,|A^{-1}||b-Ay|\,\|_\infty/\|y\|_\infty$.

Uwarunkowanie liniowego zadania najmniejszych kwadratów

Rozwiązanie $x = A^+b$ liniowego zadania najmniejszych kwadratów $Ax = b$, gdzie $A \in \mathbb{R}^{m\times n}$, $b \ne 0_m \in \mathbb{R}^m$, $m \ge n$, zaś $A^+ \in \mathbb{R}^{n\times m}$ jest macierzą pseudoodwrotną Moore'a-Penrose'a (Boullion i Odell [40], Meyer [290], Stewart [381]), w ogólności nie jest ciągłą funkcją A i w przypadku osobliwego zadania (rank $A < n$) nawet 'małe' zaburzenia A mogą powodować 'znaczne' zmiany rozwiązania. Obowiązuje bowiem oszacowanie $\|(A+\tilde{A})^+ - A^+\|_F \le 2\|\tilde{A}\|_F \max\{\|A^+\|_2^2, \|(A+\tilde{A})^+\|_2^2\}$, z którego wynika, że dla $\|\tilde{A}\|_2 \to 0$ podane górne ograniczenie nie musi dążyć do zera (Golub i Van Loan [139], Wedin [461]). Ponadto, jeżeli rank$(A+\tilde{A}) \ne$ rank A, wtedy $\|(A+\tilde{A})^+ -$

$A^+\|_2 \geq 1/\|\tilde{A}\|_2$, co oznacza możliwość nieograniczonego wzrostu normy $\|(A+\tilde{A})^+ - A^+\|_2$. Gdy $\operatorname{rank}(A+\tilde{A}) = \operatorname{rank} A$ oraz $\|A^+\|_2\|\tilde{A}\|_2 < 1$, wtedy $\|(A+\tilde{A})^+\|_2 \leq \|A^+\|_2/(1-\|A^+\|_2\|\tilde{A}\|_2)$. Na tej podstawie wnioskujemy, że gdy addytywne zaburzenie $\tilde{A}$ nie zmienia rzędu macierzy A, to taki nieograniczony wzrost nie jest możliwy (Stewart i Sun [383]). Wrażliwość zadania najmniejszych kwadratów charakteryzujemy, badając odpowiednie wskaźniki uwarunkowania (Golub *et al.* [137], Grcar [150], Wedin [461]). Dla $A \in \mathbb{R}^{m\times n}$ ($A \neq 0_{m\times n}$) definiujemy wskaźnik $\kappa_2(A) = \|A\|_2\|A^+\|_2 = \sigma_{\max}(A)/\sigma_{\min}(A)$, przy czym obowiązuje notacyjna konwencja $\sigma_{min}(A) > 0$. Zakładając nierosnące uporządkowanie wartości szczególnych macierzy A, to znaczy przyjmując, że $\sigma_1 \geq \cdots \geq \sigma_{r_A} > 0$, gdzie $r_A = \operatorname{rank} A \leq n$, otrzymujemy $\kappa_2(A) = \sigma_1/\sigma_{r_A}$.

Lemat D.5 (o ocenie względnego błędu rozwiązania liniowego zadania najmniejszych kwadratów z niestrukturalizowalnymi zaburzeniami; Higham [177], Wedin [461]). *Niech* $A, \tilde{A} \in \mathbb{R}^{m\times n}$, $m \geq n$, $\operatorname{rank} A = \operatorname{rank}(A+\tilde{A}) = n$ *(zadanie nieosobliwe) oraz:*

$$\begin{aligned} x &= \arg\min_{z\in\mathbb{R}^n} \|b - Az\|_2, \quad r = b - Ax \\ y &= \arg\min_{z\in\mathbb{R}^n} \|(b+\tilde{b}) - (A+\tilde{A})z\|_2, \quad r_y = (b+\tilde{b}) - (A+\tilde{A})y \end{aligned}$$

a ponadto $\|\tilde{A}\|_2 \leq \epsilon\|A\|_2$, $\|\tilde{b}\|_2 \leq \epsilon\|b\|_2$. *Jeżeli* $\epsilon\kappa_2(A) < 1$, *wtedy:*

$$\begin{aligned} \frac{\|x-y\|_2}{\|x\|_2} &\leq \frac{\epsilon\kappa_2(A)}{1-\epsilon\kappa_2(A)}\left(2 + (\kappa_2(A)+1)\frac{\|r\|_2}{\|A\|_2\|x\|_2}\right) \\ \frac{\|r-r_y\|_2}{\|b\|_2} &\leq \epsilon(1+2\kappa_2(A)) . \ \square \end{aligned}$$

Założenia *lematu D.5* można złagodzić przez odstąpienie od żądania pełnego kolumnowego rzędu macierzy A oraz $A+\tilde{A}$, zachowując wszelako postulat $\operatorname{rank} A = \operatorname{rank}(A+\tilde{A})$. Ponadto, gdy $\operatorname{rank} A = n$, spełnienie nierówności $\|A^+\|_2\|\tilde{A}\|_2 < 1$ wystarcza do tego, aby $\operatorname{rank}(A+\tilde{A}) = \operatorname{rank} A$. Inną ocenę błędu $\|x-y\|_2/\|x\|_2$ podano w (Demmel [84], Golub i Van Loan [139], Grcar [150], Lawson i Hanson [269] - algorytm wykorzystywany w procedurach pakietu LAPACK, Stewart i Sun [383], Van der Sluis [442], Wei [462] – algorytm możliwy do zastosowania także w przypadku osobliwych zadań najmniejszych kwadratów). Analizę względnego błędu rozwiązania nieosobliwego liniowego zadania najmniejszych kwadratów dla strukturalizowalnych zaburzeń znajdziemy w (Björck [32, 33], Higham [177], Stewart i Sun [383]). Skalowanie liniowego zadania najmniejszych kwadratów omówiono w (Golub i Van Loan [139], Higham [177], Lawson i Hanson [269]).

Istotnym krokiem przybliżającym do dostatecznie ogólnego rozwiązania problemu analizy wstecznych błędów liniowych zadań najmniejszych kwadratów były prace (Grcar [150], Gu [156], Sun [421], Waldén *et al.* [454]). W chronologicznie pierwszej pracy (Waldén *et al.* [454]) wyróżniono dwa przypadki wstecznych błędów. W pierwszym z nich dany wektor y interpretuje się jako dokładne rozwiązanie pewnego 'najbliższego' zaburzonego zgodnego układu równań. Drugi przypadek dotyczy sytuacji, gdy takie 'najbliższe' zaburzone zadanie, którego rozwiązaniem w sensie najmniejszych kwadratów jest y, nie jest zgodnym układem równań (zob. także Björck [32] oraz Higham [177]). Niech zatem $A \in \mathbb{R}^{m\times n}$ $(m \geq n)$, $b \in \mathbb{R}^m$, oraz $y \neq 0_n \in \mathbb{R}^n$. Wsteczny błąd dla niestrukturalizowalnych zaburzeń liniowego zadania najmniejszych kwadratów definiujemy jako

$$\eta_{A,b}(y) = \min\{\|[\ \tilde{A} \quad \theta\tilde{b}\ ]\|_F\ :\ y = \arg\min_{z\in\mathbb{R}^n} \|(b+\tilde{b}) - (A+\tilde{A})z\|_2\}$$

przy czym parametr θ pełni tu pomocniczą rolę: dla $\theta \to \infty$ uzyskujemy możliwość badania sytuacji, w której wektor b nie podlega zaburzeniom.

Lemat D.6 (o wstecznym błędzie rozwiązania liniowego zadania najmniejszych kwadratów dla niestrukturalizowalnych zaburzeń; Grcar [150], Higham [177], Waldén *et al.* [454]). *Niech $r = b - Ay$ będzie resztowym wektorem odpowiadającym pewnemu rozwiązaniu y nieosobliwego liniowego zadania najmniejszych kwadratów $Ax = b$. Przy założeniu niestrukturalizowalnych zaburzeń danych tego zadania, obowiązuje następująca ocena wstecznego błędu rozwiązania y*

$$\eta_{A,b}(y) = \min\{\eta_0,\ \sigma_{\min}([\ A \quad \eta_0 C\ ])\}$$

gdzie

$$\eta_0 = \sqrt{\mu}\frac{\|r\|_2}{\|y\|_2}, \qquad C = I_m - \frac{rr^T}{r^T r}, \qquad \mu = \frac{\theta^2\|y\|_2^2}{1+\theta^2\|y\|_2^2}\,. \ \square$$

Podejmowano także liczne próby wyznaczania przybliżonych ocen wstecznych błędów rozwiązań liniowych zadań najmniejszych kwadratów przy zredukowanym obliczeniowym wysiłku (Gu [156], Karlson i Waldén [213], Malyshev i Sadkane [281] - algorytm Lanczosa). Asymptotyczna ocena $\overline{\eta}_{A,b}(y) = \|(\|b-Ay\|_2^2 I_n + \|y\|_2^2 A^T A)^{-1/2} A^T(b-Ay)\|_2$ tego błędu dla przypadku, w którym: $x \neq 0_n$, $b - Ax \neq 0_m$, $y \approx x$ oraz $b - Ax \approx b - Ay$, pochodzi z (Grcar [150]).

Dla strukturalizowalnych zaburzeń korzystamy z następującej definicji wstecznego błędu liniowego zadania najmniejszych kwadratów

$$\omega_{E,f}(y) = \min\{\epsilon \, : \, y = \arg\min \|(b+\tilde{b}) - (A+\tilde{A})z\|_2 \; |\tilde{A}| \le \epsilon E, \; |\tilde{b}| \le \epsilon f\}.$$

Dla zgodnego zadania ($b \in \operatorname{Im} A$) obowiązuje formuła (A.1) ustalona w (Oettli i Prager [308]). Natomiast w przypadku, w którym $b \notin \operatorname{Im} A$, kompletne rozwiązanie problemu wstecznej analizy błędu $\omega_{E,f}(y)$ nie jest znane (Björck [32], Higham [174], Kiełbasiński i Schwetlick [219]). Przykładowo, w (Björck [32]) proponuje się oszacowanie $\omega_{|A|,|b|}(y) = \overline{\omega}_{|A|,|b|}(y)$, gdzie

$$\overline{\omega}_{|A|,|b|}(y) = \max_{1 \le i \le n} \frac{|A^T(b - Ay)|_i}{(|A^T||b - Ay|)_i}.$$

Taka ocena wstecznego błędu ma znaczenie tylko dla zadań, w których $b \notin \operatorname{Im} A$. W przypadku 'niemal' zgodnych układów równań (przykładowo dla $m = N_\Lambda$ w *algorytmie 3.1*) nie można bowiem oczekiwać, aby resztowy wektor $b - Ay$ był ortogonalny do $\operatorname{Im} A$, taki wektor zależy teraz tylko od 'numerycznego szumu', a zatem nie ma żadnej 'struktury'.

Numeryczna stabilność liniowego zadania najmniejszych kwadratów

O danej metodzie wyznaczania rozwiązania $y \in \mathbb{R}^n$ liniowego zadania najmniejszych kwadratów $Ax = b$, $A \in \mathbb{R}^{m\times n}$, mówimy, że jest numerycznie stabilna, jeżeli istnieje taka macierz $\tilde{A} \in \mathbb{R}^{m\times n}$ oraz takie wektory $\tilde{b} \in \mathbb{R}^m$ i $\tilde{y} \in \mathbb{R}^n$, że

$$\|(b+\tilde{b}) - (A+\tilde{A})(y+\tilde{y})\|_2 = \min_{z\in\mathbb{R}^n} \|(b+\tilde{b} - (A+\tilde{A})z\|_2$$

przy czym $\|\tilde{A}\|_2 \le \epsilon\|A\|_2$, $\|\tilde{b}\|_2 \le \epsilon\|b\|_2$, $\|\tilde{y}\|_2 \le \epsilon\|y\|_2$, zaś $\epsilon = \mathcal{O}(u_\varepsilon)$. Gdy spełniony jest powyższy warunek, także o rozwiązaniu y mówimy, że jest rozwiązaniem stabilnym. Innymi słowy, y jest stabilnym rozwiązaniem, jeżeli istnieją 'małe' zaburzenia $\tilde{A}$, $\tilde{b}$ oraz $\tilde{y}$, przy których $y + \tilde{y}$ jest dokładnym rozwiązaniem odpowiedniego zadania najmniejszych kwadratów. Należy pamiętać, że w przypadku źle uwarunkowanego zadania rozwiązanie y może znacząco różnić się od dokładnego rozwiązania $x = A^+b$.

Dany wektor y jest nazywany wstecznie stabilnym (*backward stable*) rozwiązaniem nieosobliwego zadania najmniejszych kwadratów, gdy istnieje taka macierz $\tilde{A} \in \mathbb{R}^{m\times n}$, dla której

$$\|b - (A+\tilde{A})y\|_2 = \min_{z\in\mathbb{R}^n} \|b - (A+\tilde{A})z\|_2$$

przy czym $\|\tilde{A}\|_2 \leq \epsilon\|A\|_2$, zaś $\epsilon = \mathcal{O}(u_\varepsilon)$ (Gu [156]). Jest oczywistym, że rozwiązanie wstecznie stabilne jest rozwiązaniem stabilnym. Minimalny (w sensie normy Frobeniusa) wsteczny błąd $\hat{\eta}_{A,b}(y)$ odpowiadający tej definicji wstecznej stabilności dany jest wzorem (Gu [156])

$$\hat{\eta}_{A,b}(y) = \begin{cases} 0 & \text{gdy } r = 0_m \\ \frac{\|A^T b\|_2}{\|b\|_2} & \text{gdy } y = 0_n \\ \min\{\eta_0, \sigma_{\min}([\ A \quad \eta_0 C\ ])\} & \text{w pozostałych przypadkach} \end{cases}$$

przy czym $r = b - Ay$, $\eta_0 = \|r\|_2/\|y\|_2$ oraz $C = I_m - rr^T/(r^T r)$. Rozwiązanie y jest zatem wstecznie stabilne, jeżeli wsteczny błąd $\hat{\eta}_{A,b}(y)$ jest dostatecznie 'mały'. Czy y może być rozwiązaniem stabilnym, gdy odpowiadający mu wsteczny błąd $\hat{\eta}_{A,b}(y)$ 'mały' nie jest? Okazuje się, że stabilne rozwiązanie y jest także rozwiązaniem wstecznie stabilnym. Oznacza to, że y jest rozwiązaniem stabilnym wtedy i tylko wtedy, gdy przyporządkowany mu wsteczny błąd $\hat{\eta}_{A,b}(y)$ jest 'mały'. Zauważmy, że C jest macierzą ortogonalnego rzutu na podprzestrzeń $\operatorname{Im} r^{\perp}$. Możliwa jest zatem sytuacja, w której $\angle\{b, \operatorname{Im} A\}$ przyjmuje 'dużą' wartość, a pomimo tego wsteczny błąd $\hat{\eta}_{A,b}(y)$ jest 'mały'. Numeryczne wyznaczanie minimalnej wartości szczególnej w przypadku macierzy o dużych wymiarach może sprawiać kłopoty (procedury xGESVD dostępne w pakiecie LAPACK wymagają $\mathcal{O}(m^3)$ arytmetycznych działań). Pewien uproszczony algorytm szacowania $\sigma_{\min}([\ A \quad \eta_0 C\ ])$ przedstawiono w (Gu [156]); algorytm ten wymaga wykonania $\mathcal{O}(mn^2)$ działań.

Sterowanie odporne polega na zapewnieniu układowi sterowania wymaganej stabilności oraz jakości w warunkach występowania niepewności w modelu sterowanego obiektu dynamicznego. Przedmiotem pracy są zagadnienia związane z syntezą liniowych algorytmów odpornego sterowania w czasie dyskretnym obiektami czasu ciągłego. Skupiono się na algorytmach wynikających z metod przestrzeni H_∞. Wskazano na znaczenie analizy uwarunkowania zadania syntezy (optymalizacji) sterowania oraz na rolę oceny numerycznych błędów proponowanych algorytmów. Omówiono sposoby polepszania uwarunkowania poprzez zastosowanie modelowania opartego na operatorze *delta*.

W pracy wykazano przydatność łańcuchowych macierzy rozproszenia modelowanego obiektu, udowodniono szereg twierdzeń odnoszących się do J-bezstratnych faktoryzacji takich macierzy, a także omówiono strukturę algorytmów sterowania optymalnych ze względu na normę H_∞.

Teoretyczne rozważania zilustrowano numerycznymi przykładami dotyczącymi zadań odpornego sterowania oraz estymacji stanu. Przykłady te obejmują między innymi: metodę rozmieszczania biegunów, sterowanie predykcyjne, a także sterowanie optymalne ze względu na kwadratowy wskaźnik jakości oraz ze względu na normę H_∞.

ISSN 0208-8029
ISBN 83-85847-94-4

W celu uzyskania bliższych informacji i zakupu dodatkowych egzemplarzy prosimy o kontakt z Instytutem Badań Systemowych PAN
ul. Newelska 6, 01-447 Warszawa
tel. 837-35-78 w. 241 e-mail: biblioteka@ibspan.waw.pl