Metody Numeryczne – Opracowanie Dodatkowe

Uwarunkowania macierzy i określanie błędów numerycznych w układach równań liniowych

Wstęp

Do przygotowania tego opracowania pobudziły mnie pytania dotyczące uwarunkowania macierzy i sposobu obliczania błędów przy rozwiązywaniu układów równań liniowych. Po przypomnieniu tych pytań i postawieniu kilku nowych udzieliłem w nim przynajmniej częściowych odpowiedzi na omawiane kwestie. Niektóre tematy (takie jak normy wektorów i macierzy) omówiłem bardziej szczegółowo w celu podkreślenia informacji, które w moim odczuciu nie są dość jasno przekazywane przez istniejące źródła.

Spis treści

Post na blogu - „Uwarunkowanie macierzy"

Poniższy fragment został pierwotnie opublikowany na moim DevBlogu pod adresem http://temporal.pr0.pl/devblog/2008/12/12/uwarunkowanie-macierzy/

Ten post jest matematyczny więc ostrzegam, że nie każdy może mieć ochotę go czytać :P. Mnie by się nie chciało

Ostatnimi czasy na studiach daje mi w kość przedmiot zwany Metodami Numerycznymi. Niedługo poprawa kolokwium, egzamin połówkowy i takie tam inne :). W związku z tym zacząłem zgłębiać temat uwarunkowania macierzy i doszedłem do dość dziwnych obserwacji.

Zacznijmy od podstaw teoretycznych, czyli jak definiujemy uwarunkowanie macierzy:

$$\mathrm{cond}(A) = ||A|| \cdot ||A^{-1}||.$$

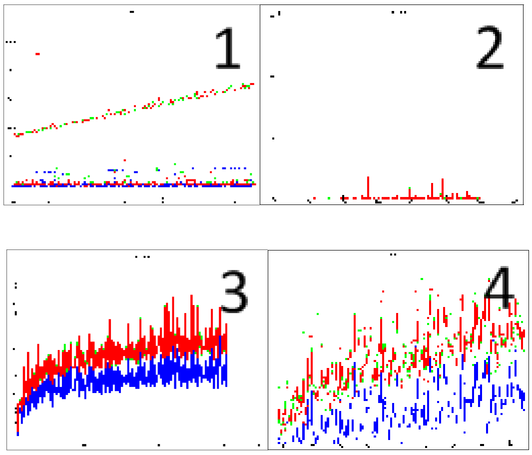
gdzie zapis ||A|| oznacza normę macierzy A. Norm możemy ustalić wiele; w przypadku macierzy stosuje się często normy: **wierszową**, **kolumnową** i **spektralną**. Nie będę tu ich tłumaczył, szkoda miejsca ;).

Wobec powyższego powstaje naturalne pytanie: *której normy użyć do obliczeń współczynnika uwarunkowania?*

W wykładach i książkach (Fortuna 🙂) niewiele znalazłem, Wikipedia też nie pomaga. Jako leniwy ale kreatywny student postanowiłem nie szukać intensywnie odpowiedzi po forach matematycznych, tylko przeprowadzić małe badania. Zadałem pytania:

- **Właściwie jak to jest z tymi normami? Czy dla danej macierzy te trzy normy są identyczne albo chociaż podobne?**
- **Jak różnią się od siebie współczynniki uwarunkowania macierzy w zależności od wybranej normy?**

W znalezieniu odpowiedzi pomógł MATLAB. Skrypt, którym przeprowadziłem badania można pobrać tutaj: [SKRYPT] (Octave-compatibile 😀). W doświadczeniu tym generowałem po 10 losowych macierzy NxN, gdzie N zmieniało się od 2 do 300. Pozwala to tak z grubsza wyrobić sobie jakąś opinię na temat tego, jak rosną normy i współczynniki uwarunkowania zależnie od rozmiaru macierzy.
Poniżej garść obrazków z MATLABa:



(powiększone wersje rysunków zamieszczone w dodatku A)

Wnioski:

- Normy wierszowe i kolumnowe (**rys. 1**, odpowiednio czerwona i zielona linia) są do siebie podobne, rosną mniej więcej liniowo. Norma spektralna (niebieska linia) również rośnie z grubsza liniowo wraz z rozmiarem macierzy, ale robi to dużo wolniej od pozostałych.
- Normy macierzy odwrotnych do macierzy A są raczej niewielkie, trzymają się blisko zera, ale rozsiane są bardzo chaotycznie. Gdzieniegdzie jakaś norma wyskoczy bardzo wysoko.
- Współczynniki uwarunkowania dla norm wierszowych i kolumnowych (**rys. 2**, kolory jak na poprzednim) są bardzo zbliżone do siebie, podobnie jak same normy. Uwarunkowanie dla

normy spektralnej jest o rząd wielkości mniejsze (**rys. 3**, to samo w skali logarytmicznej na osi y).

- Współczynniki uwarunkowania dla wszystkich norm razem rosną i maleją (**rys. 4**, tzn. wzrost współczynnika uwarunkowania dla norm wierszowej i kolumnowej wiąże się ze wzrostem tegoż współczynnika dla normy spektralnej).

Pomimo powyższych przemyśleń dalej pozostaje dla mnie nierozwiązaną jedna kwestia: **której normy używać do policzenia uwarunkowania macierzy w metodach numerycznych, kiedy chcę określić uwarunkowanie zadania?** Może mi ktoś z Czytelników pomoże :).

Postawienie problemu

Streszczając powyższy post w jednym pytaniu, można by napisać:

- **Jakiej normy użyć i dlaczego, aby wyrazić uwarunkowanie zadania?**

Aby na nie odpowiedzieć w kontekście rozwiązywania układów równań liniowych, należy wcześniej zbadać poniższe kwestie:

- **Czym charakteryzują się normy macierzy? Co je łączy, co dzieli? O czym nam mówią?**
- **Jak możemy wyznaczyć dokładność rozwiązania układu równań? Do którego miejsca po przecinku wynik jest na pewno dobry?**

Poniższa praca udziela przynajmniej częściowej odpowiedzi na te pytania.

Czym jest norma macierzy?

Przed wyjaśnieniem normy macierzy należy przypomnieć pojęcie normy wektora. Symbol $\|x\|$ oznacza *normę wektora* x. **Jest to funkcja, a jej wartością jest liczba, którą możemy interpretować jako długość wektora**. Biorąc dowolne dwa wektory x i y o takiej samej liczbie elementów, warunki, które spełnia norma wektora (za [3]) wyglądają następująco:

$$\begin{aligned}
&\|x\| \geq 0, \quad \text{dla dowolnego wektora x} \\
&\|x\| = 0 \Leftrightarrow x = (0,0,\ldots,0)^T \equiv 0 \\
&\|\alpha x\| = |\alpha| \cdot \|x\|, \quad \text{dla dowolnego wektora x i liczby rzeczywistej } \alpha \\
&\|x + y\| \leq \|x\| + \|y\|, \quad \text{dla dowolnych wektorów x i y}
\end{aligned} \tag{1}$$

Mówiąc prostym językiem norma wektora zerowego jest równa 0, norma każdego innego wektora jest dodatnia (pierwsze dwa wzory) oraz α razy dłuższy wektor ma tyle samo razy większą normę (wzór trzeci). Wzór czwarty - zwany warunkiem trójkąta - mówi nam, że droga od punktu A do C na pewno nie będzie dłuższa niż droga od punktu A do B i od punktu B do C. Wszystko to jest zgodne z intuicyjnym pojęciem długości*.

* Normy wiążą się z podobnym terminem z analizy matematycznej – z metrykami. Dokładniejsze omówienie tych drugich pod kątem ciekawych zastosowań można znaleźć w [4].

*Norma dowolnej kwadratowej** *macierzy* A oznaczana jest symbolem $\|A\|$. Dla dowolnych macierzy A i B o takiej samej liczbie elementów norma macierzy musi spełniać następujące warunki (ponownie za [3]):

$$\|A\| \geq 0$$
$$\|A\| = 0 \Leftrightarrow \quad \text{wszystkie elementy macierzy A są zerowe}$$
$$\|\alpha A\| = |\alpha| \cdot \|A\|$$
$$\|A + B\| \leq \|A\| + \|B\|$$
$$\|AB\| \leq \|A\| \cdot \|B\| \qquad (2)$$

Warunki te zgodnie z oczekiwaniami są analogiczne do tych ze wzoru (1), jednak w przypadku macierzy dodatkowo wymagamy, by norma iloczynu była mniejsza lub równa iloczynowi norm. Ostatni wzór jest szczególnie istotny dla tego opracowania.

Jeśli jednak norma wektora to po prostu jego długość, to czym jest norma macierzy? Przyjmijmy za A macierz kwadratową nxn, a za x n-elementowy wektor. *Naturalną (lub: indukowaną) normę macierzy* możemy zdefiniować następująco:

$$\|A\| = \max_{\|x\|=1} \frac{\|Ax\|}{\|x\|} \qquad (3)$$

Zauważmy, że iloczyn $v = Ax$ jest pewnym n-elementowym wektorem; można powiedzieć, że macierz A przekształca wektor x w wektor v**. **Norma macierzy wyraża zatem największą możliwą długość wektora o długości 1 po przekształceniu przez macierz A.** Warto zauważyć, że w ten sposób definiujemy normę macierzy na podstawie odpowiadającej jej normy wektora. Wszystkie normy macierzy w tym opracowaniu są normami indukowanymi.

Normy i ich zależności

Warunki (1), (2) i (3) pozwalają na ułożenie wielu różnych definicji norm. Spośród najbardziej popularnych wyróżniamy normy: wierszową, kolumnową i euklidesową (spektralną). Ich definicje dla wektorów i macierzy ilustruje poniższa tabela (na podstawie [8]):

Norma	Wektor	Macierz
$1-$norma (norma kolumnowa)	$\|x\|_1 = \sum_i \|x_i\|$	$\|A\|_1 = \max_j \sum_i \|a_{i,j}\|$
$2-$norma (norma spektralna)	$\|x\|_2 = \sqrt{\sum_i \|x_i\|^2}$	$\|A\|_2 = \max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2}$
$\infty-$norma (norma wierszowa)	$\|x\|_\infty = \max_i \|x_i\|$	$\|A\|_\infty = \max_i \sum_j \|a_{i,j}\|$

* Definicja normy macierzy obejmuje też macierze prostokątne, jednak nie spełniają one wtedy wszystkich warunków podanych we wzorze (2).

** Takie przekształcenia na macierzach 2x2, 3x3 i 4x4 stanowią istotę współczesnej grafiki komputerowej czasu rzeczywistego.

***Tabela 1:** normy wektorów i macierzy*

W szczególności normę spektralną dla macierzy można zdefiniować w inny, równoważny sposób – mianowicie jest to moduł największej co do modułu *wartości własnej*.

Wykonując badania opisane w poście zaobserwowałem związek między wszystkimi tymi normami. Okazuje się, że istnieje pojęcie *równoważności norm*. Dla dowolnych dwóch norm wektorowych $\|\cdot\|_\alpha$ i $\|\cdot\|_\beta$ można odnaleźć takie stałe, że zachodzi związek:

$$r\|A\|_\alpha \leq \|A\|_\beta \leq s\|A\|_\alpha \quad (4)$$

Przykłady takich związków (w przypadku macierzy kwadratowej), szczególnie istotne dla dokonanych obserwacji:

$$\frac{1}{\sqrt{n}}\|A\|_1 \leq \|A\|_2 \leq \sqrt{n}\|A\|_1$$

$$\frac{1}{\sqrt{n}}\|A\|_\infty \leq \|A\|_2 \leq \sqrt{n}\|A\|_\infty$$

$$\|A\|_2 \leq \sqrt{\|A\|_1 \cdot \|A\|_\infty} \quad (5)$$

Z tabeli 1 wynika, że kolumnowa i wierszowa norma macierzy może tylko rosnąć wraz ze wzrostem rozmiaru macierzy (z powodu modułów nie może maleć, a gdyby w jakimś przypadku nie wzrosła to mielibyśmy do czynienia z macierzą osobliwą, co dyskwalifikuje ją z dalszych rozważań w kontekście rozwiązywania układów równań liniowych). Zjawisko to tłumaczy wniosek z obserwacji, że norma spektralna w jakiś sposób zależy od pozostałych dwóch, gdyż rośnie razem z nimi, tylko wolniej.

Szacowanie błędów w układach równań liniowych

W obliczeniach numerycznych mogą pochodzić przede wszystkim z dwóch źródeł. Po pierwsze, obliczenia na liczbach zmiennoprzecinkowych obarczone są błędem (błąd zaokrąglenia, ang. round-off error). Po drugie, dane wejściowe nie zawsze są dokładne. Pewne liczby (np. π) nie można przechowywać na skończonej liczbie bajtów w komputerze, więc musimy operować na ich przybliżeniach. Dane mogą pochodzić z innych obliczeń albo z zewnętrznych pomiarów. Obie te kategorie błędów musimy uwzględnić przy rozwiązywaniu układów równań.

Współczynnik uwarunkowania macierzy względem danej normy definiujemy jako:

$$K(A) = \|A\| \cdot \|A\|^{-1} \quad (6)$$

Wartość tego współczynnika dla konkretnej macierzy jest zależna od wyboru normy (problem ten był tematem przedrukowanego na wstępie posta na blogu, który stał się podstawą tego opracowania).

Rozwiązując* układ równań liniowych postaci $Ax = b$ w sposób numeryczny otrzymujemy przybliżony wektor rozwiązań χ. Wstawiając go do układu w miejsce szukanego rozwiązania x (którego nie znamy!) otrzymamy wektor $\beta = A\chi$. Oznaczając wektor reszty $r = b - A\chi = b - \beta$

* Omówienie metod numerycznego rozwiązywania takich układów można znaleźć w: [1], [2], [3], [5].

możemy napisać następujące wzory (na podstawie [3]):

$$\|x-\chi\| \le K(A)\frac{\|r\|}{\|A\|} \quad (7)$$

$$\frac{\|x-\chi\|}{\|x\|} \le K(A)\frac{\|r\|}{\|b\|} \quad (8)$$

Wzory (7) i (8) informują nas, że:

- Względny błąd (wzór (8)) reszty przenosi się na względny błąd rozwiązania ze współczynnikiem K(A).
- Bezwzględny błąd (wzór (7)) reszty przenosi się na bezwzględny błąd rozwiązania ze współczynnikiem $\frac{K(A)}{\|A\|} = \|A^{-1}\|$.

Wzory te są podstawą do szacowania błędów rozwiązania. Dają nam one górne ograniczenie na wartość bezwzględną błędu. W warto w tym miejscu zastanowić się, czym tak właściwie jest $\|x-\chi\|$ oraz $\|r\|$.

- $\|x-\chi\|$ to długość różnicy między wynikiem dokładnym a jego przybliżeniem, czyli odpowiedź na pytanie: „jak daleko znajduje się wynik dokładny od mojego przybliżenia"? Należy w tym miejscu podkreślić słowo *daleko* – definicja tego terminu zależy właśnie od tego, którą normę wybierzemy. W normie euklidesowej (2-norma) przy wektorach jedno, dwu i trójelementowych słowo *daleko* odpowiada naszemu codziennemu rozumieniu odległości na prostej, płaszczyźnie i w przestrzeni. W innych normach jednak odległości definiowane są inaczej, więc słowo *daleko* znaczy po prostu co innego.
- Analogicznie $\|r\|$ to odległość między właściwym wektorem wyrazów wolnych a wektorym wyliczonym przez wstawienie do układu równań przybliżonego rozwiązania χ.

Podsumowując, wzór 7 można wyrazić słowami: „punkt* będący rozwiązaniem dokładnym jest nie dalej od mojego przybliżenia, niż $\frac{K(A)}{\|A\|}$ razy odległość między dokładnym wektorem b a tym, który uzyskałem przez wstawienie χ do układu równań". W szczególności (co trzeba koniecznie podkreślić) **nie chodzi tutaj o to, z jaką dokładnością składowe wektora χ odpowiadają wartościom dokładnym**. Ponieważ wybrana przez nas norma definiuje słowo *daleko*, zależnie od wyboru normy błędy te będą się różnić od siebie.

Na podstawie tabeli 1 można dokonać następującej interpretacji błędu bezwzględnego $\|x-\chi\|$ w zależności od doboru normy:

- $\|x-\chi\|_\infty$ - maksymalny błąd bezwzględny składowych wektora (czyli jak x_1 ma się do χ_1, jak x_2 ma się do χ_2, itd.)
- $\|x-\chi\|_2$ - wspomniana już odległość punktów x i χ w przestrzeni

* N-wymiarowy wektor „pokazuje" punkt w przestrzeni n-wymiarowej.

- $\|x-\chi\|_1$ - suma bezwzględnych błędów składowych
- $\frac{\|x-\chi\|_1}{n}$ - średni błąd bezwzględny składowej (n oznacza liczbę elementów wektora x i χ)

Szczególnie przydatna okazuje się więc $\infty -$ norma, gdyż pozwala ona oszacować maksymalny błąd wszystkich składowych.

Warto jeszcze zastanowić się, skąd właściwie biorą się wektory χ i β. Zazwyczaj zależy to od problemu. W pewnych sytuacjach uzyskujemy przybliżony wektor χ i na jego podstawie wyliczamy wektor β. Potem możemy użyć uzyskane ograniczenie błędu by spróbować innego przybliżenia rozwiązania (na tym opierają się metody iteracyjnego poprawiania, [3]). W innych możemy z góry wiedzieć, że wektor β jest obarczony błędem (na przykład pochodzi z pomiarów fizycznych lub poprzednich obliczeń) i chcemy wiedzieć jedynie jak wpłynie to na dokładność rozwiązania (czy jest sens w ogóle obliczać rozwiązania naszego układu).

Odpowiedzią na pierwsze pytanie: „Jakiej normy użyć i dlaczego, aby wyrazić uwarunkowanie zadania?" brzmi: zależy to tylko od naszego humoru, gdyż wyniki zazwyczaj są zbliżone. Należy mieć jednak na uwadze, że różne normy udzielają różnych informacji i w pewnych szczególnych przypadkach możemy potrzebować konkretnej normy. Jeden z takich przypadków omówiłem w dalszej części opracowania. **Należy jednak pamiętać, aby używać tej samej normy we wszystkich obliczeniach – zarówno błędów jak i współczynnika uwarunkowania!**

Szacowanie dokładności obliczeń

Błędy liczone w $\infty -$ normie mogą zostać użyte do oszacowania, ile cyfr w uzyskanym rozwiązaniu χ zostało obliczonych poprawnie. Wynika to z faktu, że zgodnie z definicją $\infty -$ normy wartość $\|x-\chi\|_\infty$ opisuje największy występujący w tym wektorze błąd odpowiadających sobie składowych. Jak już wspomniano, błędy te mogą pochodzić z dwóch źródeł – niedokładności obliczeń (temat ten omówiono na przykładach w [7]) i niepewności danych wejściowych w układzie równań (w tym opracowaniu pomijam problem niepewności w elementach macierzy A). Zanim przedstawię swoją metodę szacowania błędu, chciałbym wspomnieć o metodzie autorów dokumentu [6].

W [6] sugeruje się, by oszacować liczbę cyfr znaczących wyniku na podstawie błędu względnego ze wzoru (8):

$$0.5\cdot 10^{-m} \le K(A)\frac{\|r\|_\infty}{\|b\|_\infty} \quad (9)$$

m oznacza w tym miejscu liczbę *cyfr znaczących*, które są poprawne w przybliżonej wartości χ. Cyfry znaczące to cyfry wartości obarczonej błędem, które są na pewno poprawne*. Według autorów [6] takie porównanie pozwala nam określić liczbę cyfr znaczących w składowych uzyskanego wektora χ, które na pewno są poprawne.

W moim osobistym odczuciu powyższe wyliczenia nie dają nam w pełni poprawnych wyników. Liczony tutaj błąd względny niszczy nam informację o maksymalnym błędzie składowej -

* Przykładowo (na podstawie Wikipedii): 0.00102341 $\pm$ 0.000003 zmienia się w przedziale od 0.00102041 do 0.00102641, więc mamy trzy cyfry znaczące: 0.00**102**.

$\|r\|_\infty = \|b-\beta\|_\infty$ rzeczywiście pokazuje nam maksymalny błąd w składowych, ale $\frac{\|r\|_\infty}{\|b\|_\infty} = \frac{\|b-\beta\|_\infty}{\|b\|_\infty}$ już nie. Rozpatrzmy to na przykładzie:

Przyjmijmy $b = \begin{bmatrix} 2 \\ 1000 \\ 3.333 \\ 5 \end{bmatrix}$ i $\beta = \begin{bmatrix} 2 \\ 999.9911 \\ 5 \\ 5 \end{bmatrix}$. Wartość $r = \begin{bmatrix} 0 \\ 0.0089 \\ -1.667 \\ 0 \end{bmatrix}$. Odpowiednie normy to: $\|r\|_\infty = 1.667$, $\|b\|_\infty = 1000$. Przyjmując (dla uproszczenia) współczynnik uwarunkowania $K(A) = 1$ dla jednostkowej macierzy A nierówność (9) zachodzi dla $m \le 3$, czyli zgodnie z [6] co najmniej 3 cyfry znaczące w wyniku są poprawne. Czy rzeczywiście?

Rozwiązując równania $Ax = b$ i $A\chi = \beta$ otrzymujemy $x = b = \begin{bmatrix} 2 \\ 1000 \\ 3.333 \\ 5 \end{bmatrix}$ i $\chi = \beta = \begin{bmatrix} 2 \\ 999.9911 \\ 5 \\ 5 \end{bmatrix}$. Od razu widać, że w wyniku różnią się praktycznie wszystkie cyfry. Przyczyną tego zjawiska jest fakt, że największy błąd bezwzględny wystąpił dla składowych b_1 i β_1*, ale norma wektora b, przez którą dzieliliśmy, zawiera informację o składowej b_2.

Biorąc pod uwagę powyższy przykład proponuję własną metodę określenia dokładności wyniku, która korzysta z błędu bezwzględnego, a nie względnego. Skorzystajmy z wzoru (7) przyjmując macierz i wektory tak jak w przykładzie powyżej. Sprawdźmy oszacowanie:

$$0.5 \cdot 10^{-m} > K(A) \frac{\|r\|}{\|A\|} \quad (10)$$

gdzie $m+1$ oznacza liczbę miejsc po przecinku, do których możemy bezpiecznie zaokrąglić składowe wektora χ by uzyskać takie same wartości jak składowe x po zaokrągleniu do $m+1$ miejsc (tak zwane *poprawne cyfry ułamkowe*). Ujemne wartości m oznaczają zaokrąglanie do cyfr przed przecinkiem (np. -1 oznacza zaokrąglenie do setek)**. To wszystko oznacza też, że składowe nie będą się różnić o więcej niż $5 \cdot 10^m$.

Dokonując porównania na podstawie wzoru (10) otrzymamy $m \le -1$, czyli składowe wektora są poprawne po zaokrągleniu do cyfry setek i na pewno nie różnią się o więcej niż 5. Spojrzenie na przykład powyżej potwierdza te wnioski.

Osobiście wydaje mi się, że taki sposób mierzenia błędu jest przydatny w praktyce; był też to podstawowy sposób szacowania błędów rozwiązań, którego szukałem. Jest to też odpowiedź na

* Składowe n-wymiarowego wektora numeruję od 0 do n-1

** Liczba w przykładzie z przypisu na poprzedniej stronie ma 5 poprawnych cyfr ułamkowych. Liczba 155 $\pm$ 8 zmienia się w przedziale od 147 do 163, więc posiada jedną poprawną cyfrę i – w tym sprawozdaniu – -2 poprawne cyfry ułamkowe. Innymi słowy, dopiero cyfra setek jest poprawna, ale poprawne zaokrąglenie otrzymamy dopiero dla cyfry tysięcy.

pytania: „Jak możemy wyznaczyć dokładność rozwiązania układu równań? Do którego miejsca po przecinku wynik jest na pewno dobry?”.

Problem ze współczynnikiem uwarunkowania macierzy

Zgodnie ze wzorem (6) do obliczenia współczynnika uwarunkowania potrzebujemy macierzy odwrotnej A^{-1}. Jest to bardzo niepożądane – **podstawową przyczyną istnienia numerycznych metod rozwiązywania układów równań jest próba uniknięcia wyliczania tej macierzy!** Obliczenie macierzy odwrotnej jest bardzo złożone obliczeniowo - z tego powodu opracowano też metody szacowania współczynnika uwarunkowania, które nie wymagają znajomości macierzy odwrotnej A^{-1}. Jedną z takich metod, specjalnie dla metody eliminacji Gaussa, można odnaleźć w [3]. Szacuje ona współczynnik:

$$K(A) \approx \frac{\|y\|}{\|\chi\|} \cdot 10^t \qquad (11)$$

gdzie t oznacza liczbę cyfr dziesiętnych, na których prowadzimy obliczenia, a y wyliczany jest z równania $Ay = r$, a jego norma w przybliżeniu wynosi $\|y\| \approx \|x - \chi\|$. W przykładzie zamieszczonym razem z omówieniem tej metody w [3] obliczono przy pomocy 5-cio cyfrowej arytmetyki oszacowanie $K(A) \approx \frac{\|y\|}{\|\chi\|} \cdot 10^5 = 16672$ oraz wartość „dokładną” przy takiej arytmetyce, wynoszącą $K(A) = 15999$. Co ciekawe, wartość dokładna przy arytmetyce zmiennoprzecinkowej podwójnej precyzji na 64 bitach wynosi $K(A) = 14241$, czyli mniej niż oba oszacowania.

Podsumowanie – odpowiedzi na pytania

- *Jakiej normy użyć i dlaczego, aby wyrazić uwarunkowanie zadania?*

  W ogólności – która nam się bardziej podoba. Normy są równoważne. W szczególności to zależy od naszych potrzeb. Aby oszacować dokładność poszczególnych składowych wyniku należy użyć $\infty -$normy. Aby oszacować intuicyjnie rozumianą odległość wyniku od wartości dokładnej w przestrzeni n-wymiarowej należy użyć normy spektralnej. Ważne jest by używać konsekwentnie wybranej normy we wszystkich obliczeniach.

- *Czym charakteryzują się normy macierzy? Co je łączy, co dzieli? O czym nam mówią?*

  Normy macierzy spełniają te same warunki co normy wektorów. Wszystkie normy są zależne od rozmiaru macierzy. Wszystkie normy indukowane macierzy kwadratowych można traktować jako sposób wyrażenia, z jaką mocą macierz przekształca wektor.

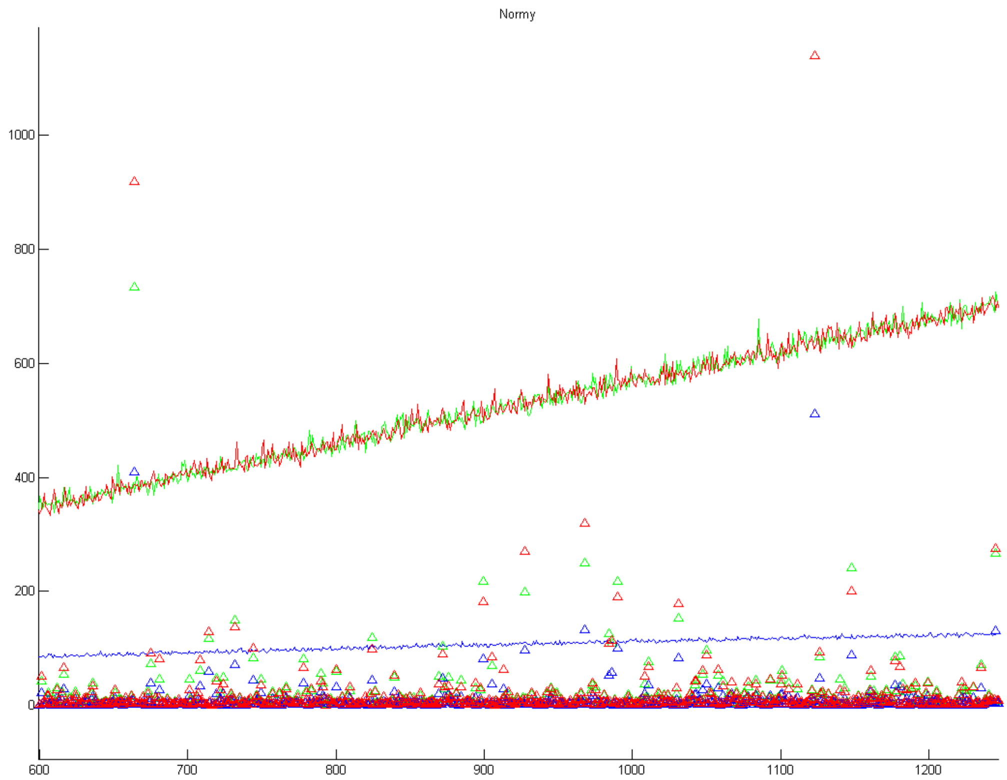
- *Jak możemy wyznaczyć dokładność rozwiązania układu równań? Do którego miejsca po przecinku wynik jest na pewno dobry?*

  Mając odchylenie wyrazów wolnych od dokładnych (znane lub uzyskane przez wstępne oszacowanie rozwiązania) oraz współczynnik uwarunkowania macierzy współczynników dla wybranej normy możemy określić górną granicę błędu normy rozwiązania. Korzystając w tym celu z $\infty -$normy możemy określić górną granicę błędu wszystkich składowych, a na jej podstawie określić, do którego miejsca po przecinku rozwiązanie jest na pewno dobre.

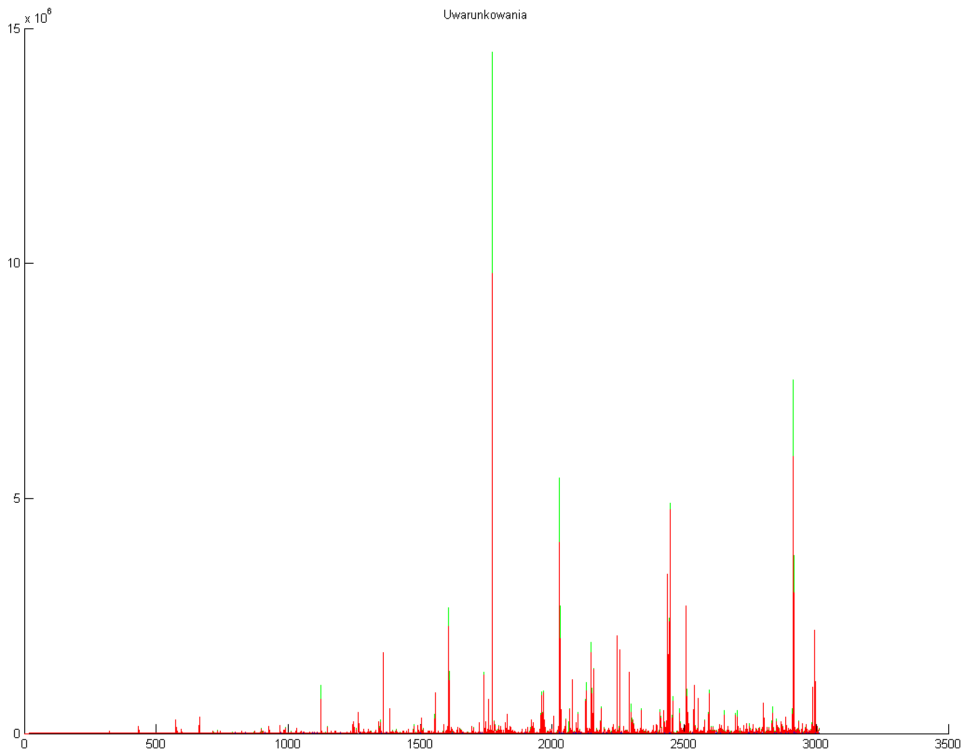
Bibliografia

1. Z. Fortuna, B. Macukow, J. Wąsowski, „Metody Numeryczne", wydanie III, wyd. WNT, Warszawa 1982, 1993
2. Åke Björck, Germund Dahlquist, „Metody Numeryczne", wydanie II, wyd. PWN, Warszawa 1987
3. R. L. Burden, J. D. Faires, „Numerical Analysis", wydanie IV, wyd. PWS-KENT Publishing Company, Boston 1989
4. DevLog Piotra Wacha, dostępny pod adresem http://wachu-dev.blogspot.com/
5. Jacek Złydach, „Metody Numeryczne – Sprawozdanie III", Kraków 2008, dostępne w Internecie
6. Dokument „Adequacy of Solutions", dostępny na stronie Holistic Numerical Methods Institute pod adresem: http://numericalmethods.eng.usf.edu/mcd/civ/04sle/mcd_civ_sle_spe_adequacy.pdf
7. Dokument „Effect of Significant Digits on Solution of Simultaneous Linear Equations", dostępny na stronie Holistic Numerical Methods Institute pod adresem: http://numericalmethods.eng.usf.edu/mcd/civ/04sle/mcd_civ_sle_spe_sigdigits.pdf
8. Fragment dokumentacji pakietu ScaLAPACK, dostępny pod adresem: http://www.netlib.org/scalapack/slug/node135.html

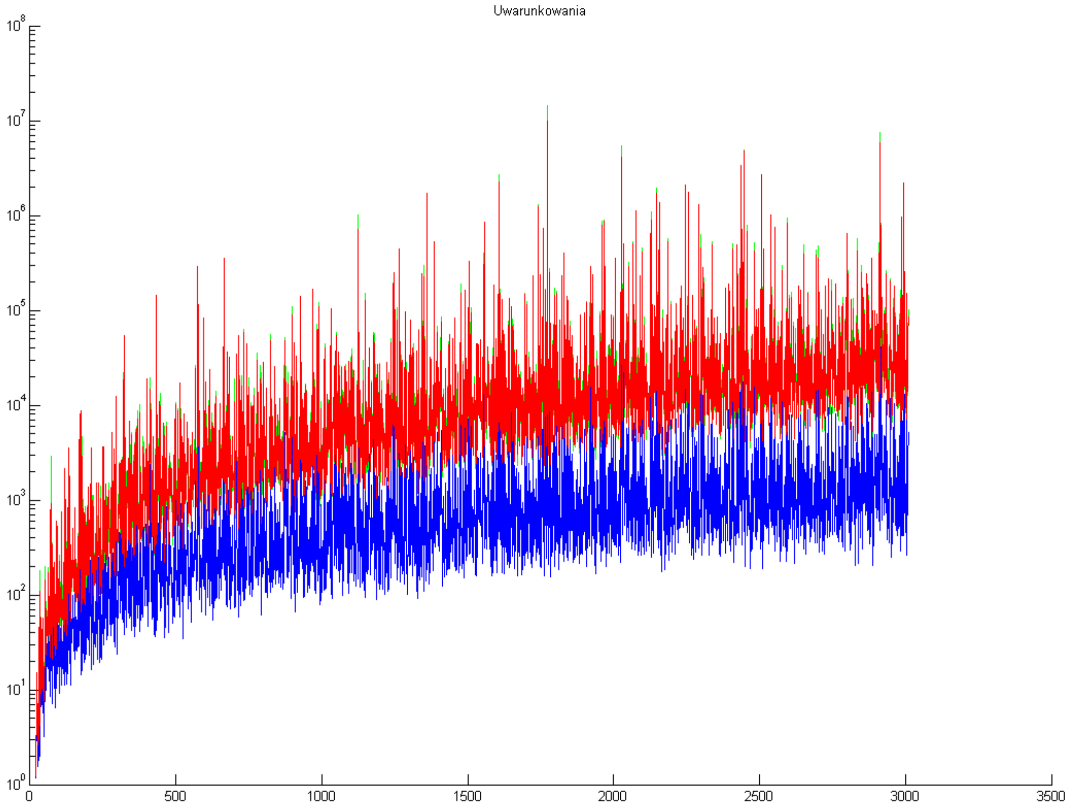
Dodatek A – kolorowe ilustracje



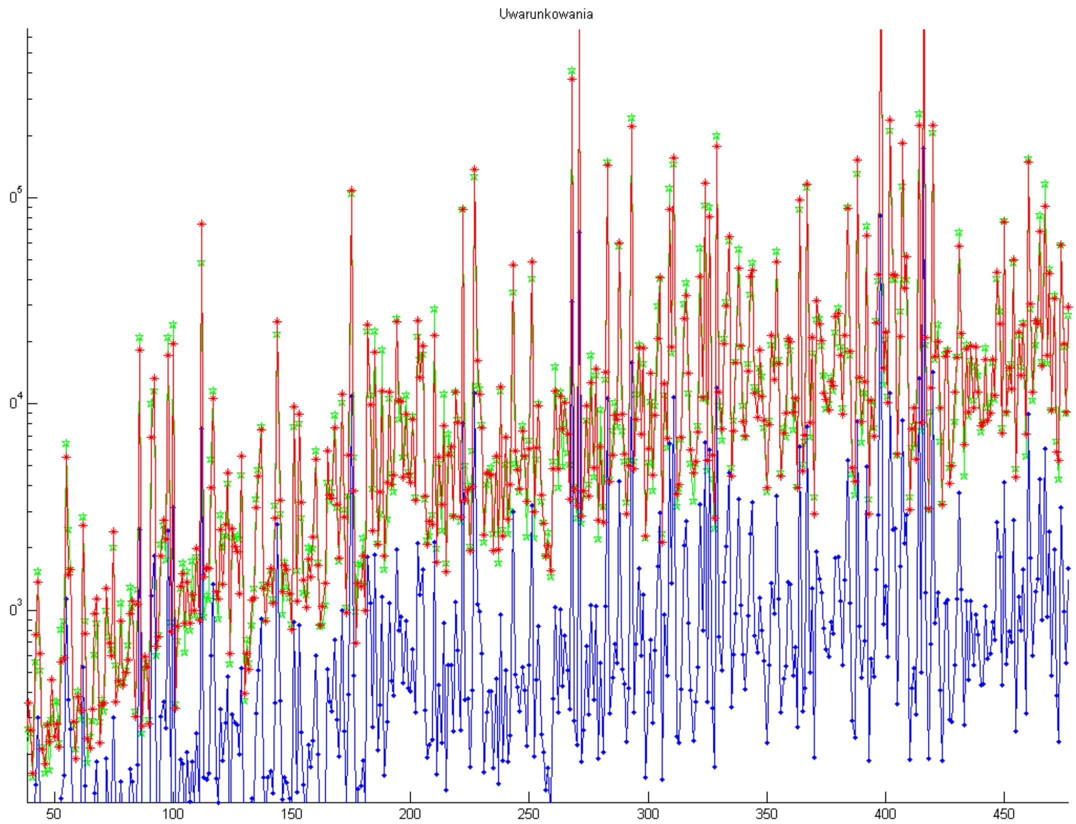
***Ilustracja 1**: Fragment wykresu norm macierzy.*



***Ilustracja 2**: Współczynniki uwarunkowania.*



***Ilustracja 3**: Współczynniki uwarunkowania - skala logarytmiczna.*



***Ilustracja 4**: Współczynniki uwarunkowania - skala logarytmiczna, zaznaczone punkty.*