Układy równań

Joanna Czarnowska[1]

[1]Uniwersytet Gdański
Instytut Informatyki

Układ równań liniowych

Układ n równań liniowych o n niewiadomych $x_1, x_2, \ldots, x_n$, współczynnikach rzeczywistych a_{ij}, $i, j = 1, 2, \ldots, n$ oraz wyrazach wolnych b_i, $i = 1, 2, \ldots, n$

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2$$

$$\ldots$$

$$a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n,$$

możemy zapisać w postaci macierzowej

$$\mathrm{A} \cdot \mathrm{x} = \mathrm{b},$$

gdzie

$$\mathrm{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad \mathrm{x} = (x_1, x_2, \ldots, x_n)^T, \quad \mathrm{b} = (b_1, b_2, \ldots, b_n)^T.$$

LU (Lower-Upper) faktoryzacja

W metodzie LU rozwiązywania układów równań liniowych, macierz $A = [a_{ij}]_{1 \leqslant i,j \leqslant n}$ zapisywana jest jako iloczyn macierzy

$$\mathrm{A} = \mathrm{L} \cdot \mathrm{U},$$

gdzie L jest dolną macierzą trójkątną, a U – górną macierzą trójkątną

$$\mathrm{L} = \begin{bmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{bmatrix}, \quad \mathrm{U} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{bmatrix}.$$

Układ równań przyjmuje wówczas postać

$$A \cdot x = \mathrm{L} \cdot \underbrace{\mathrm{U} \cdot \mathrm{x}}_{\mathrm{y}} = \mathrm{b}.$$

Jego rozwiązanie, sprowadza się do rozwiązania dwóch układów równań z macierzami trójkątnymi

$$\mathrm{L} \cdot \mathrm{y} = \mathrm{b}, \qquad \mathrm{U} \cdot \mathrm{x} = \mathrm{y}.$$

Przykład

Przyklad 1. (BF, Example 2, p.363, 404) Dla układu równań

$$\begin{aligned} x_1 - x_2 + 2x_3 - x_4 &= -8,\\ 2x_1 - 2x_2 + 3x_3 - 3x_4 &= -20,\\ x_1 + x_2 + x_3 \quad &= -2,\\ x_1 - x_2 + 4x_3 + 3x_4 &= 4, \end{aligned}$$

mamy

$$A = \begin{bmatrix} 1 & 1 & 0 & 3 \\ 2 & 1 & -1 & 1 \\ 3 & -1 & -1 & 2 \\ -1 & 2 & 3 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 4 & 1 & 0 \\ -1 & -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 3 \\ 0 & -1 & -1 & -5 \\ 0 & 0 & 3 & 13 \\ 0 & 0 & 0 & -13 \end{bmatrix} = LU.$$

Aby rozwiązać równanie

$$A\mathbf{x} = LU\mathbf{x} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 4 & 1 & 0 \\ -1 & -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 3 \\ 0 & -1 & -1 & -5 \\ 0 & 0 & 3 & 13 \\ 0 & 0 & 0 & -13 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 8 \\ 7 \\ 14 \\ -7 \end{bmatrix}$$

podstawiamy $y = U$x i rozwiązujemy równanie

$$L\mathbf{y} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 4 & 1 & 0 \\ -1 & -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 8 \\ 7 \\ 14 \\ -7 \end{bmatrix}$$

Otrzymujemy: $y_1 = 8$, $y_2 = -9$, $y_3 = 26$, $y_4 = -26$.

Przykład 1 c.d.

Następnie rozwiązujemy równanie $U\mathrm{x} = \mathrm{y}$

$$\begin{bmatrix} 1 & 1 & 0 & 3 \\ 0 & -1 & -1 & -5 \\ 0 & 0 & 3 & 13 \\ 0 & 0 & 0 & -13 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 8 \\ -9 \\ 26 \\ -26 \end{bmatrix}.$$

Ostatecznie otrzymujemy: $x_1 = 3,\ x_2 = -1,\ x_3 = 0,\ x_4 = 2.$

Zobacz Przykład 1 w skrypcie 09UkladyRownan.R

Macierz dodatnio określona

W rozwiązywaniu liniowych układów równań wykorzystuje się też rozkład Choleskiego, dla macierzy dodatnio określonych.

Definicja

Macierz $\mathsf{A} = [a_{ij}]_{1 \leqslant i,j \leqslant n}$ nazywamy dodatnio określoną, jeśli jest symetryczna oraz

$$\mathsf{x}^T \mathsf{A} x = [x_1 \ x_2 \ \dots \ x_n] \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j > 0.$$

dla każdego wektora $\mathsf{x} \neq 0$.

Macierz dodatnio określona – przykład

Przyklad 2. Macierz

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$

jest dodatnio określona.

Dla dowolnego wektora niezerowego $x = (x_1, x_2, x_3)$ mamy

$$\begin{aligned} \mathbf{x}^t A \mathbf{x} &= [x_1, x_2, x_3] \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\ &= [x_1, x_2, x_3] \begin{bmatrix} 2x_1 & - & x_2 & & \\ -x_1 & + & 2x_2 & - & x_3 \\ -x_2 & + & 2x_3 & & \end{bmatrix} \\ &= 2x_1^2 - 2x_1x_2 + 2x_2^2 - 2x_2x_3 + 2x_3^2. \end{aligned}$$

Po przekształceniu, otrzymujemy

$$x^T A x = x_1^2 + (x_1 - x_2)^2 + (x_2 - x_3)^2 + x_3^2 > 0$$

Kryterium Sylvestera

Niech A bedzie macierzą symetryczną

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

o minorach głównych

$$M_1 = a_{11}, \ M_2 = \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \ \dots, \ M_n = \det \begin{bmatrix} a_{11} & a_{12} & \dots & a_{nn} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

Kryterium Sylvestera. Macierz A jest dodatnio określona wtedy i tylko wtedy, gdy jej wiodące minory główne są dodatnie: $M_1 > 0, M_2 > 0, \ \dots \ M_n > 0$.

Przykład 2 c.d.

Mamy

$$\det A_1 = \det[2] = 2 > 0,$$

$$\det A_2 = \det \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = 4 - 1 = 3 > 0,$$

$$\det A_3 = \det \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} = 2 \det \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} - (-1) \det \begin{bmatrix} -1 & -1 \\ 0 & 2 \end{bmatrix}$$
$$= 2(4-1) + (-2+0) = 4 > 0.$$

Zobacz Przyklad 2 w skrypcie 09UkladyRownan.R

Rozkład Choleskiego

Macierz $\mathrm{A} = [a_{ij}]_{1 \leqslant i,j \leqslant n}$ dodatnio określoną, można zapisać w postaci iloczynu macierzy

$$\mathrm{A} = \mathrm{L}\mathrm{L}^T,$$

gdzie L jest dolną macierzą trójkątną, a L^T jej macierzą transponowaną macierzy L. Rozkład ten jest nazywany rozkładem Choleskiego.

Rozkład Choleskiego

Rozpisując iloczyn $A = LL^T$, otrzymujemy

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{bmatrix} \begin{bmatrix} l_{11} & l_{21} & \cdots & l_{n1} \\ 0 & l_{22} & \cdots & l_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & l_{nn} \end{bmatrix}$$

Skąd kolejno wyliczamy wyrazy macierzy L

$$\begin{aligned} a_{11} &= l_{11}^2 && \rightarrow l_{11} = \sqrt{a_{11}} \\ a_{21} &= l_{21} l_{11} && \rightarrow l_{21} = \frac{a_{21}}{l_{11}} \\ a_{22} &= l_{21}^2 + l_{22}^2 && \rightarrow l_{22} = \sqrt{a_{22} - l_{21}^2} \\ a_{32} &= l_{31} l_{21} + l_{32} l_{22} && \rightarrow l_{32} = \frac{a_{23} - l_{31} l_{21}}{l_{22}} \\ &\ldots \end{aligned}$$

$$l_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2}, \qquad l_{ji} = \frac{a_{ji} - \sum_{k=1}^{i-1} l_{jk} l_{ik}}{l_{ii}}.$$

Rozkład Choleskiego – przykład

Przyklad 3. Macierz

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 20 & 26 \\ 3 & 26 & 70 \end{bmatrix}$$

jest dodatnio określona. Rozkład Choleskiego tej macierzy

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 4 & 0 \\ 3 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$$

Zobacz Przyklad 3 w skrypcie 09UkladyRownan.R

Układ równań liniowych – faktoryzacja Choleskiego

Uwzględniając faktoryzację Choleskiego, układ równań liniowych

$$A \cdot x = b,$$

przyjmuje postać

$$L \cdot \underbrace{L^T \cdot x}_{y} = b.$$

Jego rozwiązanie sprowadza się podobnie jak w przypadku faktoryzacji LU do rozwiązania dwóch układów równań z macierzami trójkątnymi

$$L \cdot y = b, \qquad L^T \cdot x = y.$$

Generowanie z wielowymiarowego rozkładu normalnego

Rozkład Choleskiego wykorzystuje się do generowania z wielowymiarowego rozkładu normalnego.

Przyklad 4. Algorytm generowania z rozkładu $N_d(\mu, \Sigma)$.

- Wyznaczamy rozkład Choleskiego macierzy kowariancji $\Sigma = LL^T$.
- Generujemy próbę Z z rozkładu normalnego standardowego $N(0, \mathrm{I}_d)$.
- Próba $\mathrm{X} = \mu + L\mathrm{Z}$ jest próbą z rozkładu $N_d(\mu, \Sigma)$.

Zobacz Przykład 4 w skrypcie 09UkladyRownan.R

Układ równań nieliniowych

Układ równań nieliniowych ma postać

$$
\begin{aligned}
f_1(x_1, x_2, \dots, x_n) &= 0,\\
f_2(x_1, x_2, \dots, x_n) &= 0,\\
&\dots\\
f_n(x_1, x_2, \dots, x_n) &= 0,
\end{aligned}
$$

gdzie $f_1, f_2, \dots, f_n : \mathbb{R}^n \to \mathbb{R}$ są funkcjami przyporządkowującymi każdemu wektorowi $\mathrm{x} = (x_1, x_2, \dots, x_n)$ liczbę.

Układ równań nieliniowych

Jeśli zdefiniujemy odwzorowanie $F : \mathbb{R}^n \to \mathbb{R}^n$ następująco

$$F(x_1, x_2, \dots, x_n) = (f_1(x_1, x_2, \dots, x_n), f_2(x_1, x_2, \dots, x_n), \dots, f_n(x_1, x_2, \dots, x_n)), \tag{1}$$

to rozwiązanie nieliniowego układu równań, możemy sprowadzić do problemu znalezienia miejsca zerowego $x^*=(x_1^*, \dots, x_n^*)$ tego odwzorowania

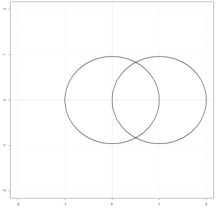
$$F(x^*) = 0.$$

Przykład. Dla układu równań

$$x_1^2 + x_2^2 - 1 = 0,$$
$$(x_1 - 1)^2 + x_2^2 - 1 = 0,$$

odwzorowanie F dane jest wzorem

$$F(x_1, x_2) = (x_1^2+x_2^2-1, (x_1-1)^2+x_2^2-1).$$



Równanie nieliniowe

Zagadnienie rozwiązania ukadu równań nieliniowych rozpoczniemy od rozwiązania równania nieliniowego, które ma postać: $f(x) = 0$, gdzie $f : \mathbb{R} \to \mathbb{R}$ jest funkcją nieliniową.

Jedną z metod rozwiązania takiego równania, jest metoda Newtona-Raphsona (metoda stycznych). Jest to algorytm iteracyjny wyznaczania przybliżonej wartości pierwiastka x^* funkcji f, na przedziale $[a, b]$.

Metoda Newtona-Raphsona

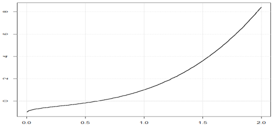
Algorytm. Niech $f \in C^2[a, b]$ (f posiada ciągłą drugą pochodną).

a) W pierwszym kroku wybieramy punkt startowy x_0 „bliski” szukanemu pierwiastkowi i taki, że $f'(x_0) \neq 0$.

b) Aproksymujemy funkcję f, wokół punktu x_0, wielomianem Taylora pierwszego rzędu
$$f(x) \approx f(x_0) + f'(x_0)(x - x_0).$$
Pierwiastek funkcji liniowej: $f(x_0) + f'(x_0)(x - x_0) = 0$
$$x = x_0 - \frac{f(x_0)}{f'(x_0)}$$
jest pierwszym przybliżeniem ($x_1 = x$) szukanej wartości x^*.

c) Dalej powtarzamy krok b), biorąc w miejsce punktu x_0 punkt x_1. Otrzymujemy
$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}.$$
Analogicznie otrzymujemy kolejne przybliżenia $(x_n)_{n\geqslant 1}$
$$x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}. \tag{2}$$

Metoda Newtona-Raphsona – przykład



Przykład 5. Rozwiązanie równania

$$x^3 + \sqrt{x} - 1 = 0$$

sprowadzamy do wyznaczenia pierwiastka funkcji

$$f(x) = x^3 + x^{1/2} - 1.$$

Wyznaczamy pochodną $f'(x) = 3x^2 + \frac{1}{2}x^{-1/2}$. Startując z punktu $x_0 = 1$, mamy

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = 1 - \frac{1}{7/2} = \frac{5}{7} = 0{,}7142857,$$

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} = 0{,}6155279,$$

$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)} = 0{,}6055141$$

$$x_4 = x_3 - \frac{f(x_3)}{f'(x_3)} = 0{,}6054234.$$

Zobacz Przykład 5a.b w 09UkladyRownan.R

Metoda Newtona-Raphsona – zbieżność

Twierdzenie. (BF: Th. 26, p.70) Jeśli funkcja f ma ciągłą drugą pochodną oraz $x^* \in (a, b)$ jest takim punktem, że $f(x^*) = 0$ i $f'(x^*) \neq 0$, to istnieje $\delta > 0$ taka, że dla każdego punktu początkowego $x_0 \in [x^* - \delta, x^* + \delta]$ metoda Newtona-Raphsona generuje ciąg punktów $(x_n)_{n=0}^{\infty}$ zbieżny do x^*.

Błąd metody. Błąd k-tego przybliżenia można oszacować poprzez nierówność

$$|x^* - x_k| \leqslant \frac{f(x_k)}{m},$$

gdzie $m = \min_{x \in [a,b]} |f'(x)|$.

Przykładowe warunki zakończenia obliczeń. Dla przyjętego ε kończy się obliczenia, jeśli:

- wartość funkcji w wyznaczonym punkcie jest bliska zeru: $|f(x_k)| \leqslant \varepsilon$,
- odległość pomiędzy kolejnymi przybliżeniami jest mała: $|x_{k+1} - x_k| \leqslant \varepsilon$,
- szacowany błąd jest dostatecznie mały: $|x^* - x_k| \leqslant \varepsilon$.

Układ równań nieliniowych – algorytm Newtona

W przypadku ukladu n równań nieliniowych z n niewiadomymi, podobnie jak dla przypadku jednowymiarowego, algorytm wyznaczania miejsca zerowego funkcji F (1), polega na wyborze wektora startowego x_0, a następnie rekurencyjnym przekształcaniu tego wektora do momentu, gdy kolejne przybliżenia będą zadawalające

$$\mathrm{x}_{n+1} = \mathrm{x}_n - J_F(\mathrm{x}_n)^{-1} F(\mathrm{x}_n),$$

gdzie J_F w powyższym równaniu macierzowym, jest macierzą Jacobiego pochodnych cząstkowych odwzorowania F

$$\begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{bmatrix}.$$

Układ równań nieliniowych – przykład

Przykład 6.

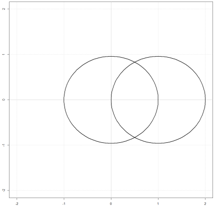
Rozwiązanie układu równań $(x_1, x_2 > 0)$

$$x_1^2 + x_2^2 - 1 = 0,$$

$$(x_1 - 1)^2 + x_2^2 - 1 = 0,$$

sprowadzamy do znalezienia miejsca zerowego odwzorowania

$$F(x_1, x_2) = (x_1^2 + x_2^2 - 1, (x_1 - 1)^2 + x_2^2 - 1).$$



Jako punkt startowy przyjmiemy $x_0 = (\frac{1}{4}, \frac{1}{2})$, dla którego: $F(x_0) = (-\frac{11}{16}, -\frac{3}{16})$.
Macierz Jacobiego w punkcie x_0 i macierz do niej odwrotna

$$J_F(x_1, x_2) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 2x_1 & 2x_2 \\ 2(x_1-1) & 2x_2 \end{bmatrix}, \quad J_F(x_0) = \begin{bmatrix} \frac{1}{2} & 1 \\ -\frac{3}{2} & 1 \end{bmatrix}, \quad J_F^{-1}(x_0) = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{3}{4} & \frac{1}{4}. \end{bmatrix}$$

Stąd

$$x_1 = x_0 - J_F(x_0)^{-1} F(x_0) = \left(\frac{1}{2}, \frac{17}{16}\right).$$

Kolejne aproksymacje zobacz Przyklad 6a w skrypcie 09UkladyRownan.R

Optymalizacja, a układy równań

Problem znalezienia rozwiązania układu równań nieliniowych, można sprowadzić też do problemu wyznaczenia punktów, dla których funkcja

$$g(x) = \sum_{i=1}^{n} f_i^2(x)$$

osiaga najmniejszą wartość (równą zero).

Przykład 6 cd. Rozwiązanie układu równań

$$x_1^2 + x_2^2 - 1 = 0,$$

$$(x_1 - 1)^2 + x_2^2 - 1 = 0,$$

sprowadzamy do wyznaczenia punktów, dla których funkcja

$$g(x_1, x_2) = (x_1^2 + x_2^2 - 1)^2 + ((x_1 - 1)^2 + x_2^2 - 1)^2$$

osiaga najmniejszą wartość (równą zero) na domkniętym przedziale.

Zobacz Przyklad 6b w skrypcie 09UkladyRownan.R