

Rozwiązywanie układów równań liniowych metodami bezpośrednimi

Plan wykładu:

1. Definicje macierzy, norm etc.
2. Metoda eliminacji Gaussa, Jordana
3. Rozkład LU metodą Gaussa
4. Układy równań z macierzą symetryczną. Rozkłady: LDL^T , LL^T
5. Układy równań z macierzą trójdziagonalną
6. Iteracyjne poprawianie rozwiązań
7. Układy liniowe nadokreślone, układ normalny równań, metoda ortogonalizacji Grama-Schmidta

Pojęcia podstawowe

Macierz jest uporządkowanym układem $m \times n$ liczb rzeczywistych lub zespolonych

$$A = (a_{ij}) \quad (i=1, \dots, m; j=1, \dots, n)$$

$$A_{m \times n} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

Jeśli $m=n$ to A jest **kwadratowa** stopnia n .

Macierz diagonalna $D = (\delta_{ij} d_{ij})$

$$D = \begin{bmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_n \end{bmatrix}$$

Macierz jednostkowa $I = (\delta_{ij})$

$$I = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

Transpozycja - $A=(a_{ij})$ to $A^T=(a_{ji})$

$$(A^T)^T = A$$

$$(\alpha A)^T = \alpha A^T$$

$$(A + B)^T = A^T + B^T$$

$$(AB)^T = B^T A^T$$

$$(Ax)^T = x^T A^T$$

Ślad macierzy $A=A_{n \times n}$

$$tr(A) = \sum_{i=1}^n a_{ii}$$

$$tr(A^T) = tr(A)$$

Macierz trójkątna lewa (dolna)

$$L = \begin{bmatrix} l_{11} & 0 & \dots & 0 \\ l_{21} & l_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ l_{n1} & l_{n2} & \dots & l_{nn} \end{bmatrix}$$

Macierz trójkątna prawa (górną)

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & \dots & r_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & r_{nn} \end{bmatrix}$$

Sumy, iloczyny i odwrotności macierzy trójkątnych tego samego rodzaju są macierzami trójkątnymi.

Wyznacznik macierzy trójkątnej

$$\det(L) = l_{11}l_{22} \dots l_{nn}$$

$$\det(R) = r_{11}r_{22} \dots r_{nn}$$

$$\det(A) = \det(A^T)$$

$$\det(AB) = \det(A)\det(B)$$

Jeśli

$$\det(A) \neq 0$$

to macierz jest **niesobliwa** i dla takiej macierzy istnieje **macierz odwrotna** A^{-1}

$$AA^{-1} = A^{-1}A = I$$

$$(AB)^{-1} = B^{-1}A^{-1}$$

Macierz symetryczna

$$A^T = A$$

Macierz ortogonalna $Q=Q_{m \times n}$ (wzajemnie ortogonalne kolumny lub wiersze)

$$Q^T Q = I$$

$$Q^{-1} Q = I$$

$$Q^T = Q^{-1}$$

Macierzą hermitowską nazywamy macierz, która po **transpozycji i sprzężeniu zespolonemu** jej elementów jest równa macierzy pierwotnej

$$A^H = (A^T)^*$$

Elementy diagonalne macierzy hermitowskiej są rzeczywiste (pozostałe mogą być zespolone lub rzeczywiste).

Macierzą dodatniookreśloną nazywamy macierz rzeczywistą lub hermitowską o własności:

$$\vec{x}^T A \vec{x} > 0, \quad \vec{x} \in R^n, \quad \vec{x} \neq 0$$

- macierz dodatnio określona jest zawsze odwracalna
- macierz odwrotna jest również dodatnio określona

Przestrzenią liniową (wektorową) nad ciałem liczb rzeczywistych (zespolonych) nazywamy zbiór obiektów (wektorów) z określonym działaniem dodawania elementów przestrzeni oraz mnożenia ich przez liczbę i oznaczamy R^N (C^N).

Aksjomaty: łączność, przemienność, element neutralny, element odwrotny,....

Uporządkowany zbiór liczb rzeczywistych (zespólonych) tworzy wektor

$$\vec{x} = (x_1, x_2, x_3, \dots, x_n)^T$$

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Przestrzeń wektorowa będąca zbiorem takich obiektów ma wymiar n.

Dowolny zbiór n wektorów **liniowo niezależnych** w \mathbb{R}^n tworzy bazę przestrzeni

$$\{\vec{y}_1, \vec{y}_2, \dots, \vec{y}_n\}$$

$$\vec{y}_\alpha \cdot \vec{y}_\beta = (\vec{y}_\alpha, \vec{y}_\beta) = (\vec{y}_\alpha)^T \vec{y}_\beta = \sum_{i=1}^n y_{\alpha i} y_{\beta i}$$

$$\mathbf{y}_\alpha \cdot \mathbf{y}_\beta = (\mathbf{y}_\alpha, \mathbf{y}_\beta) = (\mathbf{y}_\alpha)^T \mathbf{y}_\beta$$

Każdy element przestrzeni można zapisać jako kombinację liniową elementów bazy

$$\vec{x} = \alpha_1 \vec{y}_1 + \alpha_2 \vec{y}_2 + \dots + \alpha_n \vec{y}_n$$

Podprzestrzeń liniową \mathbb{R}^k w \mathbb{R}^n tworzy zbiór wszystkich wektorów

$$\vec{y}_1, \vec{y}_2, \dots, \vec{y}_k, \quad k \leq n$$

Macierz możemy traktować jako obiekt zbudowany z wektorów (wektory wierszowe lub kolumnowe).

Rzędem macierzy $A = A_{m \times n}$ $r = \text{rank}(A)$

nazywamy największą liczbę niezależnych liniowo wektorów wierszowych lub kolumnowych. Jeśli $r = m = n$ to macierz jest nieosobliwa.

Normy wektorów i macierzy

Normy wprowadza się w celu ilościowego określania własności wektorów i macierzy.

Normą wektora nazywamy funkcję, która każdemu elementowi w R^n przyporządkowuje liczbę rzeczywistą. Dla dowolnych

$$\vec{x}, \vec{y} \in R^n \quad \alpha \in R$$

norma wektora musi spełniać następujące aksjomaty:

$$\|\vec{x}\| \geq 0, \quad \|\vec{x}\| = 0 \iff \vec{x} = \vec{0}$$

$$\|\alpha\vec{x}\| = |\alpha| \cdot \|\vec{x}\|$$

$$\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$$

Dla n-wymiarowego wektora najczęściej stosowane są normy z rodziny L_p :

$$\vec{x} = [x_1, x_2, \dots, x_n]^T$$

$$\|\vec{x}\|_p = (|x_1|^p + |x_2|^p + \dots + |x_n|^p)^{\frac{1}{p}}$$

- norma pierwsza

$$p = 1, \quad \|\vec{x}\|_1 = |x_1| + |x_2| + \dots + |x_n|$$

- norma druga (euklidesowa)

$$p = 2, \quad \|\vec{x}\|_2 = (x_1^2 + x_2^2 + \dots + x_n^2)^{1/2}$$

- norma maksymalna

$$p = \infty, \quad \|\vec{x}\|_\infty = \max\{|x_1|, |x_2|, \dots, |x_n|\}$$

Dla dowolnego wektora x w przestrzeni \mathbb{R}^n prawdziwe są poniższe relacje pomiędzy normami:

$$\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\|x\|_2 \leq n\|x\|_\infty$$

Normy macierzy

Własności norm macierzy

$$\begin{aligned} \|A\| &\geq 0, \quad \|A\| = 0 \iff A = 0 \\ \|\alpha A\| &= |\alpha| \cdot \|A\| \\ \|A + B\| &\leq \|A\| + \|B\| \\ \|AB\| &\leq \|A\| \cdot \|B\| \\ \|A\vec{x}\| &\leq \|A\| \cdot \|\vec{x}\| \text{ (normy zgodne)} \end{aligned}$$

Normy zgodne - norma macierzy indukowana przez normę wektora

Macierz o m wierszach i n kolumnach można traktować jako operator liniowy przekształcający przestrzeń \mathbb{R}^m w \mathbb{R}^n .

Normę takiej macierzy można określić przy użyciu wektorów:

$$\|A\|_{pq} = \sup_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \frac{\|A\vec{x}\|_q}{\|\vec{x}\|_p}$$

gdzie: p i q oznaczają normy wektorów w przestrzeniach \mathbb{R}^n i w \mathbb{R}^m .

Mówimy, że norma $\|A\|_{pq}$ jest normą indukowaną przez normy $\|\cdot\|_p$ oraz $\|\cdot\|_q$.

Dla **$p=q$** oznaczając

$$\|A\|_p = \|A\|_{pq}$$

możemy określić następujące normy macierzy

- maksymalna suma modułów w kolumnie

$$\|A\|_1 = \max_{j=1,2,\dots,n} \sum_{i=1}^m |a_{ij}|$$

- norma spektralna

$$\|A\|_2 = \left(\max_{i=1,\dots,n} \lambda_i(AA^T) \right)^{1/2}$$

- maksymalna suma modułów w wierszu

$$\|A\|_{\infty} = \max_{i=1,2,\dots,n} \sum_{j=1}^n |a_{ij}|$$

- maksymalny moduł elementu

$$\|A\|_{1\infty} = \max_{i,j} |a_{ij}|$$

W przestrzeniach z normą $\|\cdot\|_2$ często używa się **euklidesowej (Frobeniusa)** normy macierzy:

$$\|A\|_E = \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2}$$

która nie jest indukowana żadną normą, ale spełnia ona z normą $\|\cdot\|_2$ warunek zgodności:

$$\|A\|_2 \leq \|A\|_E \|\vec{x}\|_2, \quad \vec{x} \in R^n$$

Normy macierzy mają istotne znaczenie w analizie błędów (np. błędów rozwiązania układów równań liniowych).

Rozwiązywanie układów algebraicznych równań liniowych metodami bezpośrednimi

Szukamy rozwiązania układu równań liniowych w postaci:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ \dots &= \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= b_n \end{aligned}$$

Powyższy układ równań można zapisać w postaci macierzowej:

$$A\vec{x} = \vec{b}$$

- macierz współczynników układu

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

- szukany wektor rozwiązań

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

- wektor wyrazów wolnych

$$\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Warunek rozwiązywalności układu niejednorodnego:

$$\vec{b} \in R(A)$$

$R(A)$ – podprzestrzeń liniowa rozpięta na wektorach kolumnowych macierzy A

Dla

$$R(A) = R_n$$

warunek rozwiązywalności układu jest spełniony dla każdego \mathbf{b} i rozwiązanie ma postać

$$\vec{x} = A^{-1}\vec{b}$$

Uwaga: powyższy zapis jest poprawny matematycznie,
ale numerycznie rozwiązujemy w inny sposób

- ze względu na błędy numeryczne
- brak możliwości odwrócenia macierzy (macierze o dużej liczbie wierszy/kolumn)

Układ równań z macierzą trójkątną

$$\begin{array}{rcl}
a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n & = & b_1 \\
& & a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\
& & \dots \dots \dots \dots \dots \dots \\
& & a_{nn}x_n = b_n
\end{array}$$

Zakładamy, że elementy leżące na diagonalu są niezerowe.

Rozwiązanie układu można znaleźć iteracyjnie, zaczynając od elementu x_n :

$$\begin{array}{rcl}
x_n & = & \frac{b_n}{a_{nn}} \\
x_i & = & \frac{b_i - a_{ii+1}x_{i+1} - \dots - a_{in}x_n}{a_{ii}}
\end{array}$$

W celu wyznaczenia wszystkich składowych wektora rozwiązania x należy wykonać:

- M operacji mnożenia i dzielenia

$$M = \frac{1}{2}n^2 + \frac{1}{2}n$$

- D operacji dodawania i odejmowania

$$D = \frac{1}{2}n^2 - \frac{1}{2}n$$

Uwarunkowanie rozwiązania układu równań

Wpływ błędów zaokrągleń na wynik można oszacować analizując zaburzenia danych
- macierzy **A** i wektora **b**

1) zaburzamy wektor **b**:

$$\begin{aligned} A(\vec{x} + \delta\vec{x}) &= \vec{b} + \delta\vec{b} \\ \delta\vec{x} &= A^{-1}\delta\vec{b} \\ \|\delta\vec{x}\| &\leq \|A^{-1}\| \cdot \|\delta\vec{b}\| \end{aligned}$$

2) zaburzamy elementy macierzy **A**:

$$\begin{aligned} (A + \delta A)(\vec{x} + \delta\vec{x}) &= \vec{b} \\ A\delta\vec{x} + \delta A(\vec{x} + \delta\vec{x}) &= \vec{0} \\ \delta\vec{x} &= -A^{-1}\delta A(\vec{x} + \delta\vec{x}) \\ \|\delta\vec{x}\| &\leq \|A^{-1}\| \cdot \|\delta A\| \cdot \|\vec{x} + \delta\vec{x}\| \end{aligned}$$

ostatnią nierówność zapisujemy w postaci

$$\frac{\|\delta\vec{x}\|}{\|\vec{x} + \delta\vec{x}\|} \leq \|A^{-1}\| \cdot \|A\| \frac{\|\delta A\|}{\|A\|}$$

$$\frac{\|\delta\vec{x}\|}{\|\vec{x} + \delta\vec{x}\|} \leq \kappa(A) \frac{\|\delta A\|}{\|A\|}$$

gdzie

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

jest **wskaźnikiem uwarunkowania macierzy**

Korzystając z wyniku (1) oraz nierówności

$$\|\vec{b}\| = \|A\vec{x}\| \leq \|A\| \cdot \|\vec{x}\|$$

dostajemy oszacowanie na błąd względny rozwiązania

$$\frac{\|\delta\vec{x}\|}{\|\vec{x}\|} \leq \kappa(A) \frac{\|\delta\vec{b}\|}{\|\vec{b}\|}$$

Wniosek - duży wskaźnik uwarunkowania macierzy może powodować duże względne zaburzenia rozwiązania nawet dla małych zaburzeń wektora danych. Zadanie jest wówczas **źle uwarunkowane**.

Metoda eliminacji Gaussa rozwiązania układu równań liniowych

Metoda jest dwuetapowa: (1) przekształcamy macierz do postaci trójkątnej,
 (2) rozwiązujemy układ z macierzą trójkątną

Etap 1: eliminacja zmiennych

Układ pierwotny

$$A^{(1)}\vec{x} = \vec{b}^{(1)}$$

$$\begin{aligned} a_{11}^{(1)}x_1 + a_{12}^{(1)}x_2 + \dots + a_{1n}^{(1)}x_n &= b_1^{(1)} \\ a_{21}^{(1)}x_1 + a_{22}^{(1)}x_2 + \dots + a_{2n}^{(1)}x_n &= b_2^{(1)} \\ \dots &= \dots \\ a_{n1}^{(1)}x_1 + a_{n2}^{(1)}x_2 + \dots + a_{nn}^{(1)}x_n &= b_n^{(1)} \end{aligned}$$

Odejmujemy od i-tego wiersza ($i=2,3,\dots,n$) wiersz pierwszy pomnożony przez współczynnik

$$l_{i1} = \frac{a_{i1}^{(1)}}{a_{11}^{(1)}}$$

Z równań $i=2,3,\dots,n$ wyeliminowana została zmienna x_1

$$A^{(2)}\vec{x} = \vec{b}^{(2)}$$

$$\begin{aligned}
 a_{11}^{(2)} x_1 + a_{12}^{(2)} x_2 + \dots + a_{1n}^{(2)} x_n &= b_1^{(2)} \\
 a_{22}^{(2)} x_2 + \dots + a_{2n}^{(2)} x_n &= b_2^{(2)} \\
 \dots\dots\dots &= \dots \\
 a_{n2}^{(2)} x_2 + \dots + a_{nn}^{(2)} x_n &= b_n^{(2)}
 \end{aligned}$$

Powtarzamy operację, ale odejmujemy od i-tego wiersza ($i=3,4,\dots,n$) wiersz drugi pomnożony przez współczynnik

$$l_{i2} = \frac{a_{i2}^{(2)}}{a_{22}^{(2)}}$$

Postępując dalej w ten sposób eliminujemy z każdego następnego równania jedną zmienną. Eliminację kończymy po $(n-1)$ krokach, gdy uzyskamy trójkątny układ równań w postaci:

$$\begin{aligned}
 a_{11}^{(n)} x_1 + a_{12}^{(n)} x_2 + \dots + a_{1n}^{(n)} x_n &= b_1^{(n)} \\
 a_{22}^{(n)} x_2 + \dots + a_{2n}^{(n)} x_n &= b_2^{(n)} \\
 \dots\dots\dots &= \dots \\
 a_{nn}^{(n)} x_n &= b_n^{(n)}
 \end{aligned}
 \quad \longrightarrow \quad A^{(n)} \vec{x} = \vec{b}^{(n)}$$

Etap drugi - postępowanie odwrotne

Rozwiązanie (kolejne składowe wektora x) znajdujemy stosując wzór iteracyjny dla macierzy trójkątnej.

Wyznaczenie rozwiązania metodą Gaussa (**przekształcenie macierzy do postaci trójkątnej**) wymaga wykonania:

- M operacji mnożenia i dzielenia

$$M = \frac{1}{3}n^3 + n^2 - \frac{1}{3}n$$

- D operacji dodawania i odejmowania

$$D = \frac{1}{3}n^3 + \frac{1}{2}n^2 - \frac{5}{6}n$$

Problem niestabilności metody

Metoda eliminacji w tej postaci jest niestabilna numerycznie
– problem dzielenia przez 0 lub liczbę bliską zeru.

Rozwiązanie problemu niestabilności:

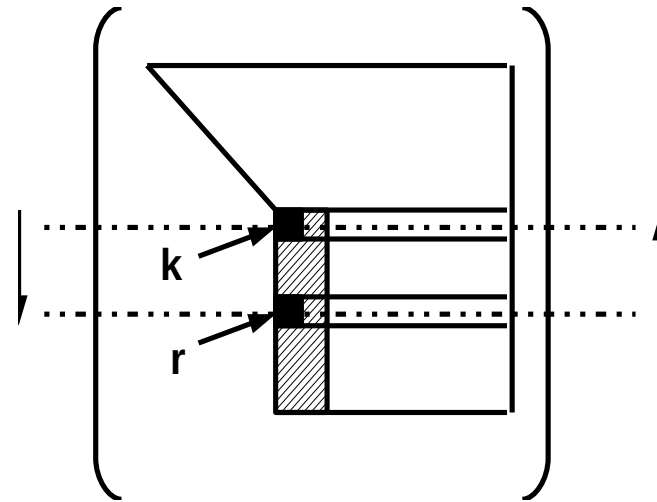
- częściowy wybór elementów głównych
- pełny wybór elementów głównych

Częściowy wybór elementów głównych

W k-tym kroku szukamy elementu

$$|a_{rk}^{(k)}| = \max_{k \leq i \leq n} |a_{ik}^k|$$

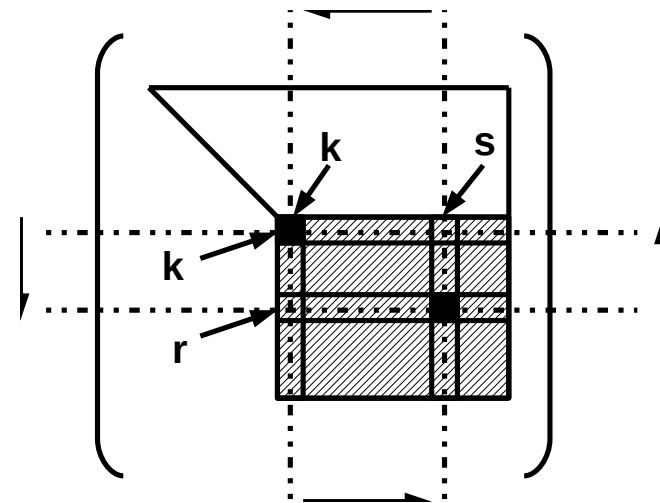
i przestawiamy wiersze r oraz k.

**Pełny wybór elementów głównych**

W k-tym kroku szukamy elementu

$$|a_{rs}^{(k)}| = \max_{k \leq i, j \leq n} |a_{ij}^k|$$

i przestawiamy wiersze: k i r oraz kolumny: k i s



- stosując wybór elementu głównego rozwiązanie otrzymujemy zawsze
- w trakcie wyboru elementu głównego należy zmienić także kolejność w \mathbf{x} i \mathbf{b}
- modyfikacji tej można nie stosować dla:
 - macierzy z dominującą przekątną

$$|a_{ii}| \geq \sum_{j=1, j \neq i}^n |a_{i,j}| \quad (i = 1, \dots, n)$$

- macierzy symetrycznej i jednocześnie dodatniookreślonej

Metoda eliminacji Jordana (eliminacji zupełnej)

W układzie równań

$$A^{(1)}\vec{x} = \vec{b}^{(1)} \quad \longrightarrow \quad \begin{array}{rcl} a_{11}^{(1)}x_1 + a_{12}^{(1)}x_2 + \dots + a_{1n}^{(1)}x_n & = & b_1^{(1)} \\ a_{21}^{(1)}x_1 + a_{22}^{(1)}x_2 + \dots + a_{2n}^{(1)}x_n & = & b_2^{(1)} \\ \dots\dots\dots & = & \dots \\ a_{n1}^{(1)}x_1 + a_{n2}^{(1)}x_2 + \dots + a_{nn}^{(1)}x_n & = & b_n^{(1)} \end{array}$$

równanie pierwsze dzielimy obustronnie przez współczynnik

$$w_1 = a_{11}^{(1)}$$

Następnie odejmujemy od **i-tego wiersza** ($i=2,3,\dots,n$) wiersz pierwszy przemnożony przez

$$w_{1i} = a_{i1}^{(1)}$$

i otrzymujemy

$$A^{(2)}\vec{x} = \vec{b}^{(2)} \quad \longrightarrow \quad \begin{array}{rcl} x_1 + a_{12}^{(2)}x_2 + \dots + a_{1n}^{(2)}x_n & = & b_1^{(2)} \\ a_{22}^{(2)}x_2 + \dots + a_{2n}^{(2)}x_n & = & b_2^{(2)} \\ \dots\dots\dots & = & \dots \\ a_{n2}^{(2)}x_2 + \dots + a_{nn}^{(2)}x_n & = & b_n^{(2)} \end{array}$$

Podobnie postępujemy z równaniem drugim, dzielimy je przez

$$w_2 = a_{22}^{(2)}$$

Następnie od i-tego wiersza ($i=1,3,4,\dots,n$) odejmujemy wiersz drugi pomnożony przez współczynnik

$$w_{2i} = a_{i2}^{(2)}$$

Otrzymujemy zmodyfikowany układ równań

$$A^{(3)}\vec{x} = \vec{b}^{(3)} \quad \longrightarrow \quad \begin{array}{rcl} x_1 + 0 \cdot x_2 + a_{13}^{(3)} x_3 + \dots + a_{1n}^{(3)} x_n & = & b_1^{(3)} \\ x_2 + a_{23}^{(3)} x_3 + \dots + a_{2n}^{(3)} x_n & = & b_2^{(3)} \\ \dots & = & \dots \\ a_{n3}^{(3)} x_3 + \dots + a_{nn}^{(3)} x_n & = & b_n^{(3)} \end{array}$$

Po przeprowadzeniu ($n-1$) eliminacji zmiennych otrzymujemy rozwiązanie

$$\begin{array}{rcl} x_1 & = & b_1^{(n)} \\ x_2 & = & b_2^{(n)} \\ \dots & = & \dots \\ x_n & = & b_n^{(n)} \end{array}$$

Liczba operacji

$$M = \frac{1}{2}n^3 + \frac{1}{2}n^2 \quad D = \frac{1}{2}n^3 - \frac{1}{2}n^2$$

Rozkład LU metodą Gaussa

Metodę Gaussa można użyć do znalezienia takich macierzy **L** i **U**, które z macierzą **A** związane są relacją:

$$A = L \cdot U$$

Procedura wyznaczania elementów tych macierzy nosi nazwę **rozkładu LU**.

Sposób postępowania (wykorzystujemy **metodę eliminacji Gaussa**)

1) mnożenie wiersza pierwszego przez czynnik

$$l_{i1} = \frac{a_{i1}^{(1)}}{a_{11}^{(1)}}$$

i odjęcie go od i-tego wiersza ($i=2\dots n$), zastępujemy mnożeniem przez macierz:

$$L^{(1)} = \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ -l_{21} & 1 & 0 & \dots & 0 \\ -l_{31} & 0 & 1 & 0 & 0 \\ \dots & \dots & \dots & 1 & 0 \\ -l_{n1} & 0 & 0 & \dots & 1 \end{bmatrix}_{n \times n}$$

co można zapisać macierzowo:

$$L^{(1)} A^{(1)} = A^{(2)}$$

$$L^{(1)} \vec{b}^{(1)} = \vec{b}^{(2)}$$

Eliminacja zmiennej z równań ($i=3,4,\dots,n$) wygląda podobnie. Mnożymy wiersze zmodyfikowanego układu równań o indeksach $i=3,4,\dots,n$ przez czynnik

$$l_{i2} = \frac{a_{i2}^{(2)}}{a_2^{(2)}}$$

i odejmujemy od nich wiersz drugi.

Operację tą można przeprowadzić mnożąc układ równań obustronnie przez macierz $L^{(2)}$:

$$L^{(2)} = \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & 0 & 0 & \dots \\ 0 & -l_{32} & 1 & 0 & 0 \\ \dots & \dots & \dots & 1 & 0 \\ 0 & -l_{n2} & 0 & 0 & 1 \end{bmatrix}$$

Zapis macierzowy operacji

$$\begin{aligned} L^{(2)} A^{(2)} &= A^{(3)} \\ L^{(2)} \vec{b}^{(2)} &= \vec{b}^{(3)} \end{aligned}$$

Po wykonaniu ($n-1$) takich operacji dostajemy

$$L^{(n-1)} L^{(n-2)} \dots L^{(1)} A^{(1)} = A^{(n)}$$

$$L^{(n-1)} L^{(n-2)} \dots L^{(1)} \vec{b}^{(1)} = \vec{b}^{(n)}$$

Macierze $L^{(i)}$ są nieosobliwe - można znaleźć dla każdej macierz odwrotną.

Przemnażając obie strony powyższych równań przez $(L^{(n-1)})^{-1}, (L^{(n-2)})^{-1}, \dots$, otrzymamy

$$\begin{aligned} A^{(1)} &= \left(L^{(1)}\right)^{-1} \left(L^{(2)}\right)^{-1} \dots \left(L^{(n-1)}\right)^{-1} A^{(n)} \\ \vec{b}^{(1)} &= \left(L^{(1)}\right)^{-1} \left(L^{(2)}\right)^{-1} \dots \left(L^{(n-1)}\right)^{-1} \vec{b}^{(n)} \end{aligned}$$

wprowadzamy oznaczenia

$$\begin{aligned} L &= \left(L^{(1)}\right)^{-1} \left(L^{(2)}\right)^{-1} \dots \left(L^{(n-1)}\right)^{-1} \\ U &= A^{(n)} = \left(L^{(n-1)} L^{(n-2)} \dots L^{(1)}\right) A^{(1)} \\ A &= L \cdot U \end{aligned}$$

Jak znaleźć macierze $(L^{(i)})^{-1}$? $L^{(i)} \left(L^{(i)}\right)^{-1} = I$

$$L^{(1)} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ -l_{21} & 1 & \dots & 0 \\ \dots & \dots & 1 & 0 \\ -l_{n1} & 0 & \dots & 1 \end{bmatrix} \quad \left(L^{(1)}\right)^{-1} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ l_{21} & 1 & \dots & 0 \\ \dots & \dots & 1 & 0 \\ l_{n1} & 0 & \dots & 1 \end{bmatrix}$$

macierz L jest macierzą dolną z jedynkami na diagonalu

$$L = (L^{(1)})^{-1} (L^{(2)})^{-1} \dots (L^{(n-1)})^{-1}$$

$$L = \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ l_{21} & 1 & 0 & \dots & 0 \\ l_{31} & l_{32} & 1 & 0 & 0 \\ \dots & \dots & \dots & 1 & 0 \\ l_{n1} & l_{n2} & l_{n3} & \dots & 1 \end{bmatrix}$$

macierz U jest macierzą górną z niezerowymi elementami na diagonalu

$$U = \begin{bmatrix} u_{11} & u_{12} & u_{13} & \dots & u_{1n} \\ 0 & u_{22} & u_{23} & \dots & u_{2n} \\ 0 & 0 & u_{33} & \dots & u_{nn} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & u_{nn} \end{bmatrix}$$

Dysponując macierzami **L** i **U** można rozwiązać układ równań

$$A\vec{x} = \vec{b}$$

$$LU\vec{x} = \vec{b}$$

$$L \underbrace{(U\vec{x})}_{\vec{y}} = \vec{b}$$

poprzez rozwiązanie 2 układów równań

$$L\vec{y} = \vec{b}$$

$$U\vec{x} = \vec{y}$$

- rozwiązanie każdego z równań wiąże się z nakładem obliczeń jak dla układu z macierzą trójkątną ($\sim 1n^2$)
- rozkład LU (eliminacja Gaussa) to nakład rzędu $\sim n^3/2$

Zalety stosowania rozkładu LU:

- Duża wydajność dla dużej liczby równań. Rozkład LU opłaca się stosować w przypadku rozwiązywania wielu układów równań z tą samą macierzą współczynników układu A. Każdy układ równań różni się wtedy tylko wektorem wyrazów wolnych. Rozkład LU wykonuje się w takim przypadku tylko raz (ilość operacji $\sim n^3$).
- Oszczędność zajmowanej pamięci. Elementy macierzy L i U mogą zostać zapisane w macierzy A.
- Jeśli macierz A jest symetryczna i dodatniookreślona to nie trzeba dokonywać wyboru elementów podstawowych.

Inne zastosowania rozkładu LU

Obliczanie wyznacznika

Aby obliczyć wyznacznik macierzy A możemy posłużyć się rozkładem

$$A = LU$$

$$\det(A) = \det(LU) = \underbrace{\det(L)}_{=1} \det(U) = \det(U) = \prod_{i=1}^n u_{i,i}$$

Wyznacznik macierzy U jest iloczynem elementów stojących na diagonalu tej macierzy (n-1 operacji mnożenia).

Odwracanie macierzy

Aby znaleźć przy pomocy macierzy L i U macierz odwrotną A^{-1} należy rozwiązać n układów równań

$$LU\mathbf{x}^{(i)} = \mathbf{e}^{(i)}, \quad i = 1, 2, \dots, n$$

wersor w przestrzeni wektorowej \mathbb{R}^n

$$\mathbf{e}^{(i)} = [0, 0, \dots, 1, \dots, 0]^T$$

$$LU \underbrace{[\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(n)}]}_X = \underbrace{[\mathbf{e}^{(1)}, \mathbf{e}^{(2)}, \dots, \mathbf{e}^{(n)}]}_I \quad \longrightarrow \quad LUX = I \rightarrow X = A^{-1}$$

Przykład Znaleźć macierz A^{-1} dla macierzy

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

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{3} & -\frac{1}{2} & 1 \end{bmatrix} \quad U = \begin{bmatrix} 3 & 3 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{bmatrix}$$

$$x^{(1)} = \begin{bmatrix} 1/6 \\ 1/6 \\ -1/6 \end{bmatrix} \quad x^{(2)} = \begin{bmatrix} -1/4 \\ 1/4 \\ 1/4 \end{bmatrix} \quad x^{(3)} = \begin{bmatrix} 1/2 \\ -1/2 \\ 1/2 \end{bmatrix} \quad P_n = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$$

$$A^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{6} & -\frac{1}{4} \\ -\frac{1}{2} & \frac{1}{6} & \frac{1}{4} \\ \frac{1}{2} & -\frac{1}{6} & \frac{1}{4} \end{bmatrix}$$

Szacowanie dokładności rozwiązania

Definiujemy **wektor reszt**

$$\vec{r} = \vec{b} - A\vec{x}$$

dla dokładnego rozwiązania mielibyśmy

$$\vec{r} = \vec{0}$$

ale rozwiązanie numeryczne zawsze różni się od dokładnego (**z powodu błędów numerycznych**)

$$\vec{r} = \vec{b} - A\vec{\tilde{x}} = \vec{b} - A(\vec{x} + \delta\vec{x}) = \underbrace{(\vec{b} - A\vec{x})}_{=0} - A\delta\vec{x} \neq \vec{0}$$

Miarą jakości rozwiązania jest norma wektora reszt

$$\|\vec{r}\| > 0$$

Dostajemy niezerowy wektor reszt - co możemy z tym zrobić?

Iteracyjne poprawianie rozwiązania układu równań

Błąd rozwiązania można sprawdzić obliczając **wektor reszt**:

$$\mathbf{r} = \mathbf{b} - A\tilde{\mathbf{x}} \quad \tilde{\mathbf{x}} - \text{numeryczne}$$

Zazwyczaj współrzędne wektora \mathbf{r} są różne od zera.
Oznacza to, że nie uzyskaliśmy dokładnego rozwiązania, ale przybliżone.

Rozwiązanie to chcemy poprawić

$$\tilde{\mathbf{x}} = \mathbf{x} - \delta\mathbf{x} \quad \mathbf{x} - \text{dokładne}$$

gdzie: $\delta\mathbf{x}$ jest poprawką, którą można łatwo wyznaczyć rozwiązując układ:

$$A\delta\mathbf{x} = \mathbf{r}$$

Należy jednak pamiętać, że wyznaczona poprawka do rozwiązania również jest przybliżeniem.

Kolejne **poprawione rozwiązanie**, które uzyskamy będzie miało postać

$$\bar{\mathbf{x}} = \tilde{\mathbf{x}} + \delta\mathbf{x} + \delta(\delta\mathbf{x})$$

Zazwyczaj procedurę tą stosuje się 2-3 krotnie, większa liczba iteracji nie poprawia rozwiązania ze względu na duży wkład błędów numerycznych wnoszony do kolejnych poprawek.

Układy równań z macierzą symetryczną

Rozkład LDL^T (macierz może nie być dodatniookreślona)

Oznaczmy rozkład **LU** dla odróżnienia jako $A = L\bar{U}$

Szukamy rozkładu macierzy A w postaci

$$A = LDU$$

gdzie: L - macierz trójkątna dolna z jedynkami na diagonalu, D - macierz diagonalna z elementami diagonalnymi macierzy \bar{U} , U - macierz trójkątna górna z jedynkami na diagonalu

Wykorzystujemy symetrię macierzy

$$U^T D L^T = A^T = A \Rightarrow U = L^T$$

co prowadzi do rozkładu dla macierzy symetrycznych

$$A = LDL^T$$

Rozwiązanie układu $A\vec{x} = \vec{b}$

Rozwiązujemy kolejno 3 układy

$$LD \underbrace{L^T \vec{x}}_{\vec{y}} = L \underbrace{D\vec{y}}_{\vec{z}} = L\vec{z} = \vec{b} \quad \longrightarrow \quad \begin{aligned} L\vec{z} &= \vec{b} \\ D\vec{y} &= \vec{z} \\ L^T \vec{x} &= \vec{y} \end{aligned}$$

Elementy rozkładu wyznaczamy rekurencyjnie

$$d_1 = a_{11}$$

a dla $i=2,3,\dots,n$ oblicza się na przemian

$$l_{ij} = \frac{a_{ij} - \sum_{k=1}^{j-1} c_{ik}l_{jk}}{d_j}$$

$$c_{ij} = d_j l_{ij}, \quad j = 1, 2, \dots, i-1$$

$$d_i = a_{ii} - \sum_{k=1}^{i-1} c_{ik}l_{ik}$$

Nakład obliczeń

$$M = \frac{1}{6}n^3 + n^2 - \frac{7}{6}n \quad D = \frac{1}{6}n^3 - \frac{1}{6}n \quad - \text{dwukrotnie mniej niż dla rozkładu LU, dzięki symetrii macierzy}$$

Zalety:

- nakład obliczeń dwukrotnie mniejszy niż dla rozkładu LU
- dzięki symetrii macierzy wystarczy zapamiętać ~połowę elementów macierzy

$$N = \frac{n(n+1)}{2}$$

Rozkład LL^T (Banachiewicza-Cholesky'ego)

Jeśli macierz A jest macierzą **symetryczną dodatniookreśloną** wówczas można znaleźć następujący rozkład

$$A = LL^T$$

Macierz L jest macierzą trójkątną dolną z elementami na diagonalu mogącymi się różnić od 1.
Macierz

$$\tilde{L} = -L$$

spełnia warunek

$$A = \tilde{L}\tilde{L}^T$$

więc rozkład ten nie jest jednoznaczny.

Jeśli jednak liczby na diagonalu macierzy L są dodatnie wówczas rozkład jest jednoznaczny, a elementy macierzy wyznaczamy ze wzorów

$$l_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2}, \quad i = 1, 2, \dots, n$$

$$l_{ji} = \frac{a_{ji} - \sum_{k=1}^{i-1} l_{jk}l_{ik}}{l_{ii}}, \quad j = i + 1, i + 2, \dots, n$$

Nakład obliczeń dla rozkładu LL^T

$$n - \sqrt{\quad}$$

$$M = \frac{1}{6}n^3 + \frac{1}{2}n^2 - \frac{2}{3}n$$

$$D = \frac{1}{6}n^3 + \frac{1}{2}n^2 + \frac{1}{3}n$$

- dwukrotnie mniej niż dla rozkładu **LU**

Przykład

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

$$i = 1 : \quad l_{11} = 2, \quad l_{21} = 1, \quad l_{31} = 1$$

$$i = 2 : \quad l_{22} = 2, \quad l_{32} = 1$$

$$i = 3 : \quad l_{33} = 2$$

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

Układy równań z macierzą trójdziagonalną

Szukamy rozwiązania układu równań

$$Tx = b$$

Zdarza się, że macierz układu równań ma postać (np. równania z ilorazami różnicowymi)

$$T = \begin{bmatrix} d_1 & c_1 & & & & \\ a_2 & d_2 & c_2 & & & \\ & a_3 & d_3 & c_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & c_{n-1} \\ & & & & a_n & d_n \end{bmatrix}$$

Można wykonać rozkład LU macierzy T, macierze te mają postać dwu-diagonalną

$$L = \begin{bmatrix} 1 & & & & & \\ l_2 & 1 & & & & 0 \\ & l_3 & 1 & & & \\ & & \ddots & \ddots & & \\ 0 & & & \ddots & \ddots & \\ & & & & l_n & 1 \end{bmatrix} \quad U = \begin{bmatrix} u_1 & c_1 & & & & \\ & u_2 & c_2 & & & 0 \\ & & u_3 & c_3 & & \\ & & & \ddots & \ddots & \\ & & & & \ddots & \\ 0 & & & & & c_{n-1} \\ & & & & & u_n \end{bmatrix} \quad 38$$

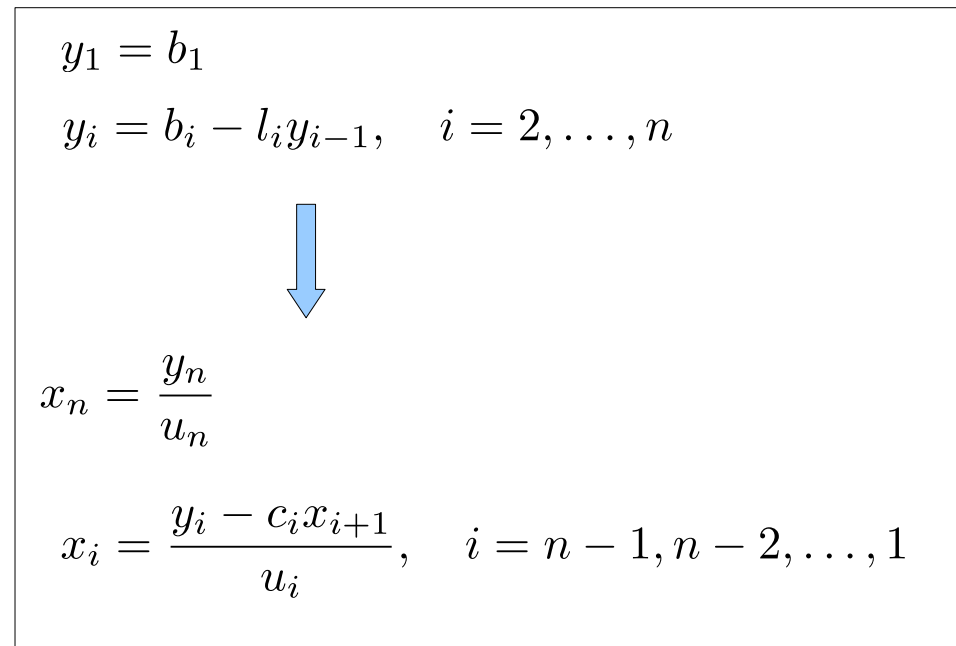
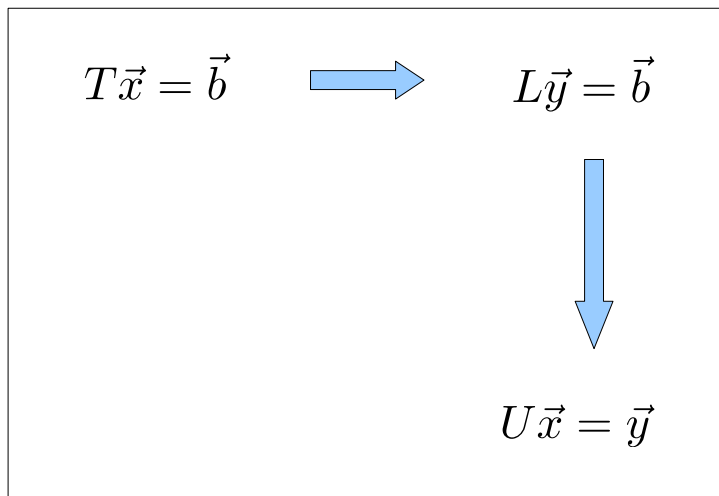
Elementy macierze rozkładu obliczamy iteracyjnie

$$u_1 = d_1$$

$$l_i = \frac{a_i}{u_{i-1}}$$

$$u_i = d_i - l_i c_{i-1}, \quad i = 2, 3, \dots, n$$

Ogólne rozwiązanie układu jest dwuetapowe jak dla **LU**



Uwagi:

- nakład obliczeń: $M=2n-2$, $D=n-1$
- liczba zajętych komórek: **$P=3n-2$**
- jeśli macierz jest **dominująca kolumnowo** to rozkład $T=LU$ jest równoważny rozkładowi z częściowym wyborem elementu podstawowego (niezawodność metody).

Rozwiązywanie układów liniowych nadokreślonych (minimalizacja formy kwadratowej)

Jak rozwiązać poniższy problem?

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ \dots \\ \dots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ \dots \\ \dots \\ b_m \end{bmatrix}$$

$$A \in R^{m \times n} \quad m > n$$

$$\mathbf{x} \in R^n$$

$$\mathbf{b} \in R^m$$

- brak dokładnego rozwiązania w większości przypadków
- można poszukiwać conajwyżej „**najlepszego przybliżenia**” rozwiązania w sensie średniokwadratowym

Dla

$$\mathbf{r} = \mathbf{b} - A\mathbf{x}$$

rozwiązaniem **średniokwadratowym** problemu nadokreślonego (**least square problem**) jest taki wektor \mathbf{x} , który minimalizuje normę

$$\|\mathbf{r}\|_2 = (\mathbf{r}^T \mathbf{r})^{1/2}$$

Szukamy minimum funkcjonału F:

$$(\mathbf{r}^T \mathbf{r})^{1/2} = \min \rightarrow \mathbf{r}^T \mathbf{r} = F = \min$$

- pozbywamy się pierwiastka, będzie łatwiej

$$\begin{aligned} F &= (\mathbf{b} - A\mathbf{x})^T (\mathbf{b} - A\mathbf{x}) \\ &= (\mathbf{b}^T - \mathbf{x}^T A^T) (\mathbf{b} - A\mathbf{x}) \\ &= \mathbf{b}^T \mathbf{b} - \mathbf{b}^T A\mathbf{x} - \mathbf{x}^T A^T \mathbf{b} + \mathbf{x}^T A^T A\mathbf{x} \end{aligned}$$

Minimum znajdziemy jeśli narzucimy warunek (różniczkujemy po kolejnych elementach wektora)

$$\frac{\partial F}{\partial \vec{x}} = \vec{0}$$

$$\frac{\partial F}{\partial \mathbf{x}} = -\mathbf{b}^T A - A^T \mathbf{b} + A^T A\mathbf{x} + \mathbf{x}^T A^T A = 0$$

Dwa pierwsze wyrazy są identyczne, (podobnie 3 i 4) więc otrzymujemy warunek

$$A^T A\mathbf{x} - A^T \mathbf{b} = 0$$

$$A^T A\mathbf{x} = A^T \mathbf{b} \quad (A^T A)^{-1} \cdot /$$

$$\mathbf{x} = (A^T A)^{-1} A^T \mathbf{b} \quad - \text{rozwiązanie istnieje}$$

$$\begin{aligned} (\mathbf{b}^T A)_j &= \sum_i b_i a_{i,j} \\ &= \sum_i a_{j,i} b_i \\ &= (A^T \mathbf{b})_j \end{aligned}$$

Jeśli macierz \mathbf{A} jest macierzą o rozmiarach $m \times n$ i elementach rzeczywistych, \mathbf{b} jest wektorem m -elementowym, a \mathbf{x} wektorem n elementowym spełniającym równanie, to układ nadokreślony możemy przekształcić do **układu normalnego**

$$A^T A \mathbf{x} = A^T \mathbf{b} \quad - \text{układ normalny}$$

$$\mathbf{z}^T \cdot / \quad A^T (\mathbf{b} - A \mathbf{x}) = \vec{0} \quad - \text{wektor } \mathbf{z} \text{ dowolny}$$

$$\underbrace{(A \mathbf{z})^T}_{\mathbf{y}^T} \underbrace{(\mathbf{b} - A \mathbf{x})}_r = 0 \quad (= \text{liczba})$$

$$\mathbf{y}^T \mathbf{r} = 0$$

wnioski:

- wektor reszt jest ortogonalny do dowolnego wektora będącego kombinacją liniową wektorów kolumnowych A (**bo te nie tworzą bazy zupełnej**)
- nie można zredukować elementów wektora reszt do zera
- zawsze dostaniemy rozwiązanie przybliżone i nie jest to wynik błędów numerycznych, ale samej konstrukcji układu równań i własności macierzy

Przykład - wpływ uwarunkowania macierzy na rozwiązanie układu normalnego

$$A = \begin{bmatrix} 1 & 1 & 1 \\ \varepsilon & 0 & 0 \\ 0 & \varepsilon & 0 \\ 0 & 0 & \varepsilon \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, |\varepsilon| \ll 1$$

Przekształcamy do układu normalnego

$$A^T A = \begin{bmatrix} 1 + \varepsilon^2 & 1 & 1 \\ 1 & 1 + \varepsilon^2 & 1 \\ 1 & 1 & 1 + \varepsilon^2 \end{bmatrix}$$
$$A^T b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

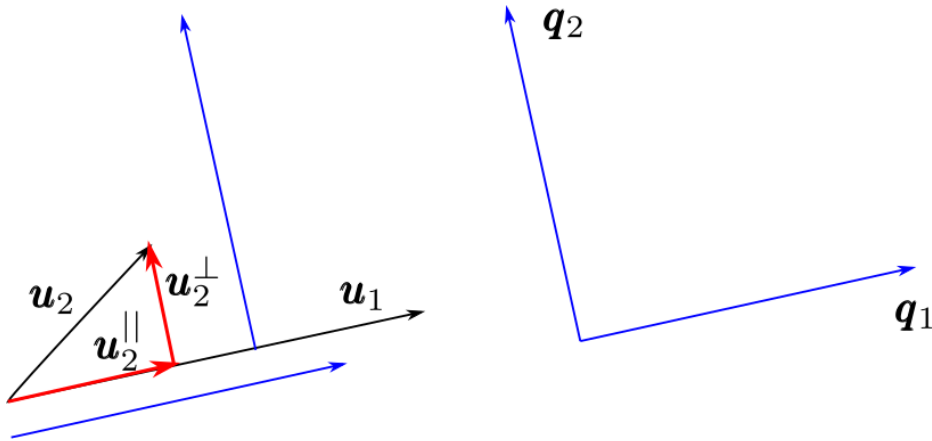
Dla precyzji obliczeń rzędu ε

$$\varepsilon^2 \approx 0$$

i macierz $A^T A$ staje się osobliwa - układ nie posiada rozwiązania.

Ortogonalizacja Grama-Schmidta

Interpretacja geometryczna



Dane są 2 wektory liniowo niezależne
i nieortogonalne: u_1, u_2

Należy je przekształcić tak, aby były ortogonalne
i unormowane

$$\{u_1, u_2\} \implies \{q_1, q_2\}$$

$$u_i^T u_j \neq 0 \quad q_i^T q_j = \delta_{i,j}$$

Jako q_1 przyjmujemy kierunek u_1 a wektor normujemy $q_1 = \frac{u_1}{\|u_1\|}$

Dla u_2 z rysunku odczytujemy

$$u_2 = u_2^\perp + u_2^\parallel \quad \longleftrightarrow \quad u_2 = u_2^\perp + r q_1 \quad q_1^T \cdot /$$

$$q_1^T u_2 = \underbrace{q_1^T u_2^\perp}_{=0} + r \underbrace{q_1^T q_1}_{=1} \quad \implies \quad r = q_1^T u_2$$

$$u_2^\perp = u_2 - q_1 (q_1^T u_2)$$

Jako drugi wektor (unormowany) przyjmujemy

$$q_2 = \frac{u_2^\perp}{\|u_2^\perp\|}$$

Uogólnienie dla k-elementowej bazy w n-wymiarowej przestrzeni

$$\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k\}, \quad \mathbf{u}_i \in R^n$$

Szukamy bazy wektorów **ortonormalnych** (ortogonalnych i unormowanych)

Jako pierwszy wektor wybieramy jak poprzednio

$$\mathbf{q}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}$$

kolejne wektory ortogonalizujemy do elementów już znalezionych

$$\tilde{\mathbf{q}}_j = \mathbf{u}_j - \sum_{i=1}^{j-1} \mathbf{q}_i (\mathbf{q}_i^T \mathbf{u}_j), \quad j = 2, 3, \dots, k$$

i normalizujemy

$$\mathbf{q}_j = \frac{\tilde{\mathbf{q}}_j}{\|\tilde{\mathbf{q}}_j\|}$$

Wektory \mathbf{q}_j możemy użyć jako np. kolumn macierzy Q

$$Q = Q_{n \times k} = [\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_k]$$

$$Q^T Q = D = \text{diag}\{d_1, d_2, \dots, d_k\} = I$$

Metody wykorzystujące rozkład QR

Dla macierzy A o rozmiarach $m \times n$, w której kolumny są niezależne liniowo istnieje jednoznaczny rozkład w postaci

$$A = QR$$

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ a_{31} & a_{32} & \dots & a_{3n} \\ \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} q_{11} & q_{12} & \dots & q_{1n} \\ q_{21} & q_{22} & \dots & q_{2n} \\ q_{31} & q_{32} & \dots & q_{3n} \\ \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots \\ q_{m1} & q_{m2} & \dots & q_{mn} \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & \dots & r_{2n} \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & r_{nn} \end{bmatrix}$$

- Q jest macierzą ortogonalną

$$Q^T Q = D$$

- D jest macierzą diagonalną o elementach nieujemnych

$$D = \text{diag}(d_1, d_2, \dots, d_n)$$

$$d_k > 0, \quad k = 1, \dots, n$$

R jest macierzą trójkątną górną z elementami

$$r_{kk} = 1, \quad k = 1, \dots, n$$

Warunek minimalizacji normy wektora reszt w sensie średniokwadratowym przyjmuje postać

$$A\mathbf{x} = \mathbf{b} \Rightarrow A^T A\mathbf{x} = A^T \mathbf{b}$$

$$R^T Q^T Q R \mathbf{x} = R^T Q^T \mathbf{b}$$

$$R^T D R \mathbf{x} = R^T Q^T \mathbf{b}$$

$$D R \mathbf{x} = Q^T \mathbf{b}$$

$$R \mathbf{x} = D^{-1} Q^T \mathbf{b} = \mathbf{y}$$

- wiemy jak rozwiązać układ z macierzą trójkątną

Jak wyznaczyć macierze Q i R? - użyjemy ortogonalizacji Gama-Schmidta

Wyznaczanie rozkładu QR metodą Grama-Schmidta

Wyznaczamy ciąg macierzy (macierz A przekształcamy iteracyjnie)

$$A = A^{(0)} \rightarrow A^{(1)} \rightarrow \dots \rightarrow A^{(n)} = Q$$

$$A^{(k)} = \left(\mathbf{q}_1, \dots, \mathbf{q}_{k-1}, \mathbf{a}_k^{(k)}, \dots, \mathbf{a}_n^{(k)} \right)$$

$$\mathbf{q}_i = \begin{bmatrix} q_{1,i} \\ \dots \\ q_{m,i} \end{bmatrix} \quad \mathbf{a}_i^{(k)} = \begin{bmatrix} a_{1,i}^{(k)} \\ \dots \\ a_{m,i}^{(k)} \end{bmatrix}$$

kolumny ortogonalne: $\mathbf{q}_1, \dots, \mathbf{q}_{k-1}$

kolumny nieprzekształcone: $\mathbf{a}_k^{(k)}, \dots, \mathbf{a}_n^{(k)}$

W k-tej iteracji ortogonalizujemy k-tą kolumnę jak w metodzie G-S

$$\mathbf{q}_k = \mathbf{a}_k - \sum_{i=1}^{k-1} \mathbf{q}_i r_{ik}$$

$$r_{ik} = \frac{\mathbf{q}_i^T \mathbf{a}_k}{d_i}, \quad i = 1, 2, \dots, k-1$$

$$\mathbf{q}_k = \mathbf{q}_k^{(k)}, \quad d_k = \mathbf{q}_k^T \mathbf{q}_k, \quad r_{kk} = 1$$

Uwaga:
wektorów \mathbf{q}_i nie normujemy
(oznaczałoby to stratę informacji o długości
pierwotnych wektorów bazowych \mathbf{a}_i)