

Metody numeryczne

Faktoryzacje

P. F. Góra

http://https://zfs.fais.uj.edu.pl/pawel_gora

2022

Kiedy eliminacja Gaussa nie wystarcza

Eliminacji Gaussa warto używać do rozwiązywania pojedynczego równania

$$Ax = b, \quad \det A \neq 0 \quad (1)$$

lub do kilku takich równań, o ile ich wszystkie prawe strony są z góry znane. (Przykładem takiej sytuacji jest rozwiązywanie liniowych równań macierzowych, w tym **niezwykle egzotyczna** sytuacja, w której należy znaleźć jawną odwrotność danej macierzy.) Widzimy wszakaże, że kolejne kroki eliminacji Gaussa zależą wyłącznie od elementów macierzy A — prawe strony, choć trzeba je przekształcać, są niejako “biernymi” uczestnikami

procesu. Należy pomyśleć o algorytmach, które zajmują się samą macierzą, odkładając przekształcanie prawych stron (kolumn wyrazów wolnych) do czasu, gdy będzie to *naprawdę* potrzebne.

Takimi algorytmami są *faktoryzacje*, czyli przedstawienie macierzy A jako iloczynu dwu macierzy w jakimś sensie prostszych:

$$A = Y \cdot Z. \quad (2)$$

Faktoryzacja LU

Przypuśćmy, że udało nam się znaleźć faktoryzację

$$\mathbf{A} = \mathbf{L} \cdot \mathbf{U}, \quad (3)$$

gdzie macierz \mathbf{U} jest trójkątna górna (wszystkie elementy poniżej głównej przekątnej są zerami), natomiast \mathbf{L} jest trójkątna dolna; dodatkowo przyjmujemy, że jej wszystkie elementy diagonalne są równe 1, $l_{ii} = 1$. Taką faktoryzację nazywamy *faktoryzacją LU*.

Jeżeli faktoryzacja LU jest znana, równanie

$$\mathbf{Ax} \equiv \mathbf{L} \underbrace{\mathbf{Ux}}_{\mathbf{y}} = \mathbf{b} \quad (4)$$

rozwiązujemy jako

$$\mathbf{Ly} = \mathbf{b} \quad (5a)$$

$$\mathbf{Ux} = \mathbf{y} \quad (5b)$$

Pierwsze z tych równań rozwiązujemy metodą *forward substitution*, drugie — metodą *back substitution*. Ponieważ są to równania z macierzami trójkątnymi, koszt obliczeniowy rozwiązania każdego z nich wynosi $O(N^2)$, a zatem koszt rozwiązania (4) wynosi $O(2N^2)$.

Pozostaje jeszcze “tylko” dokonać samej faktoryzacji.

Algorytm Doolittle'a

Aby dokonać faktoryzacji LU , należy obliczyć N^2 nieznanych elementów macierzy L , U . Rozpiszmy (3):

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1N} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2N} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3N} \\ \dots & \dots & \dots & \dots & \dots \\ a_{N1} & a_{N2} & a_{N3} & \dots & a_{NN} \end{bmatrix}}_A = \underbrace{\begin{bmatrix} 1 & & & & \\ l_{21} & 1 & & & \\ l_{31} & l_{32} & 1 & & \\ \dots & \dots & \dots & \dots & \\ l_{N1} & l_{N2} & l_{N3} & \dots & 1 \end{bmatrix}}_L \underbrace{\begin{bmatrix} u_{11} & u_{12} & u_{13} & \dots & u_{1N} \\ & u_{22} & u_{23} & \dots & u_{2N} \\ & & u_{33} & \dots & u_{3N} \\ & & & \dots & \dots \\ & & & & u_{NN} \end{bmatrix}}_U \quad (6)$$

Okazuje się, że rozwiązywanie równań na poszczególne elementy l_{ij} , u_{pq} jest proste, jeżeli przeprowadza się je *we właściwej kolejności*, odpowiadającej kolejnym kolumnom macierzy A .

Pierwsza kolumna: Aby znaleźć pierwszą kolumnę macierzy A , mnożymy kolejne wiersze L przez pierwszą kolumnę macierzy U . **Ale ta kolumna ma tylko jeden element.** Otrzymujemy

$$\begin{aligned} u_{11} &= a_{11} \\ l_{21}u_{11} &= a_{21} \\ l_{31}u_{11} &= a_{31} \\ \dots &\dots \dots \\ l_{N1}u_{11} &= a_{N1} \end{aligned} \tag{7}$$

Z pierwszego z równań (7) obliczamy u_{11} , a następnie z kolejnych $l_{21}, l_{31}, \dots, l_{N1}$.

Druga kolumna: Wyrażenia na elementy drugiej kolumny macierzy \mathbf{A} powstają z przemnożenia kolejnych wierszy \mathbf{L} przez drugą kolumnę \mathbf{U} :

$$\begin{array}{rcccccc}
 & & & & u_{12} & = & a_{12} \\
 & & & & l_{21}u_{12} & + & u_{22} & = & a_{22} \\
 & & & & l_{31}u_{12} & + & l_{32}u_{22} & = & a_{32} \\
 & & & & \dots & \dots & \dots & \dots & \dots \\
 & & & & l_{N1}u_{12} & + & l_{N2}u_{22} & = & a_{N2}
 \end{array} \tag{8}$$

Z pierwszego z tych równań obliczamy u_{12} . W tym momencie u_{12} jest już znane, podobnie jak obliczone wcześniej $l_{\bullet 1}$, a zatem z drugiego z równań (8) obliczamy u_{22} , a z kolejnych $l_{32}, l_{42}, \dots, l_{N2}$.

I tak dalej.

Widać, że średni koszt obliczenia któregoś z nieznanymi elementami l_{ij} , u_{pq} jest rzędu $O(N)$. Ponieważ elementów tych jest N^2 , złożoność numeryczna algorytmu Doolittle'a wynosi $O(N^3)$. Całkowity koszt rozwiązania układu równań liniowych, a więc faktoryzacji LU i rozwiązania układów równań z macierzami trójkątnymi (5), jest taki sam, jak eliminacji Gaussa.

Przewaga faktoryzacji LU nad eliminacją Gaussa polega na tym, iż **przy pomocy faktoryzacji LU można rozwiązywać dowolnie wiele równań z takimi samymi lewymi stronami (macierzami)**, przy czym “kosztowną” część, a więc samą faktoryzację, oblicza się tylko raz.

Z uwagi na symetrię problemu i na **kolejność wykonywanych obliczeń**, faktoryzacja LU nie wymaga dodatkowej pamięci do zapamiętania obliczonych elementów faktoryzacji: elementy macierzy L (bez diagonal) zapamiętujemy w poddiagonalnym trójkącie macierzy A , elementy macierzy U — na diagonal i w ponaddiagonalnym trójkącie A .

Przykład

W celu dokonania faktoryzacji LU macierzy

$$\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix} \quad (9)$$

musimy rozwiązać równania

$$\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ l_{21} & 1 & \\ l_{31} & l_{32} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ & u_{22} & u_{23} \\ & & u_{33} \end{bmatrix} \quad (10)$$

ze względu na l_{ik} , u_{kj} . W tym celu zapiszmy indywidualne równania, na jakie rozpada się (10), w kolejności odpowiadające przeglądaniu macierzy (9) kolumnami.

Pierwsza kolumna macierzy (9) odpowiada

$$u_{11} = 1 \quad (11a)$$

$$l_{21}u_{11} = 2 \quad (11b)$$

$$l_{31}u_{11} = 2 \quad (11c)$$

skąd natychmiast otrzymujemy

$$u_{11} = 1, \quad l_{21} = 2, \quad l_{31} = 2. \quad (12)$$

Zwróćmy uwagę, iż pierwsze z równań (11) służy do wyliczenia elementu macierzy U , drugie i trzecie — do wyliczenia elementów macierzy L .

Druga kolumna odpowiada

$$u_{12} = 2 \quad (13a)$$

$$l_{21}u_{12} + u_{22} = 1 \quad (13b)$$

$$l_{31}u_{12} + l_{32}u_{22} = 2 \quad (13c)$$

Zauważmy, że jeśli równania (13) rozwiązywać w kolejności „naturalnej”, od góry do dołu, każde z nich okazuje się być równaniem z *jedną* niewiadomą. Pierwsze dwa służą do wyliczenia elementów macierzy U , trzecie do wyliczenia elementu macierzy L . Otrzymujemy

$$u_{12} = 2, \quad u_{22} = -3, \quad l_{32} = \frac{2}{3}. \quad (14)$$

Trzecia kolumna (9) daje

$$u_{13} = 2 \quad (15a)$$

$$l_{21}u_{13} + u_{23} = 2 \quad (15b)$$

$$l_{31}u_{13} + l_{32}u_{23} + u_{33} = 1 \quad (15c)$$

W tym wypadku wszystkie trzy równania (15) służą do obliczenia elementów macierzy U . Podobnie jak poprzednio, jeśli równania te rozwiązywać

od góry do dołu, każde z nich jest równaniem z jedną niewiadomą. Jako rozwiązanie otrzymujemy

$$u_{13} = 2, \quad u_{23} = -2, \quad u_{33} = -\frac{5}{3}. \quad (16)$$

Ostatecznie

$$\begin{bmatrix} 1 & & & \\ 2 & 1 & & \\ 2 & \frac{2}{3} & 1 & \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 \\ & -3 & -2 \\ & & -\frac{5}{3} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}. \quad (17)$$

Równość w (17) można sprawdzić bezpośrednim rachunkiem.

Algorytm Crouta

Przedstawiony algorytm nie zawiera wyboru elementu podstawowego (pivotingu), ten zaś jest niezbędny dla stabilności całego procesu. Z uwagi na symetrię faktoryzacji, tylko częściowy wybór elementu podstawowego jest możliwy. Omówimy to na przykładzie. Rozwiązując równania (8) począwszy od drugiego z nich, obliczamy

$$\begin{aligned}l_{22}u_{22} &= a_{22} - l_{21}u_{12} && (l_{22} \equiv 1) \\l_{32}u_{22} &= a_{32} - l_{31}u_{12} \\&\dots \\l_{N2}u_{22} &= a_{N2} - l_{N1}u_{12}\end{aligned}\tag{18}$$

Porównujemy teraz wyliczone lewe strony równań (18) i wybieramy największą (na moduł) z nich; tę uznajemy za “właściwe” u_{22} — odpowiada to permutacji wierszy macierzy A . **Należy także spermutować już obliczone wiersze macierzy L .** W rezultacie otrzymujemy faktoryzację LU nie

samej macierzy A , ale macierzy różniącej się od niej pewną permutacją wierszy.

Przykład

Rozpatrzmy problem znalezienia następującej faktoryzacji:

$$\begin{bmatrix} 2 & 4 & 1 & 1 \\ 1 & 2 & 3 & 1 \\ 0 & 1 & 2 & -1 \\ -1 & 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & l_{32} & 1 & 0 \\ l_{41} & l_{42} & l_{43} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ 0 & u_{22} & u_{23} & u_{24} \\ 0 & 0 & u_{33} & u_{34} \\ 0 & 0 & 0 & u_{44} \end{bmatrix}. \quad (19)$$

Faktoryzację znajdujemy przechodząc macierz A *kolumnami*, poczynając od lewego górnego rogu. Pierwsza kolumna daje zatem

$$a_{11} : \quad u_{11} = 2 \quad (20a)$$

$$a_{21} : \quad l_{21}u_{11} = 1 \quad (20b)$$

$$a_{31} : \quad l_{31}u_{11} = 0 \quad (20c)$$

$$a_{41} : \quad l_{41}u_{11} = -1 \quad (20d)$$

Po przejściu pierwszej kolumny faktoryzacja ma postać

$$\begin{bmatrix} 2 & 4 & 1 & 1 \\ 1 & 2 & 3 & 1 \\ 0 & 1 & 2 & -1 \\ -1 & 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & 1 & 0 & 0 \\ 0 & l_{32} & 1 & 0 \\ -\frac{1}{2} & l_{42} & l_{43} & 1 \end{bmatrix} \begin{bmatrix} 2 & u_{12} & u_{13} & u_{14} \\ 0 & u_{22} & u_{23} & u_{24} \\ 0 & 0 & u_{33} & u_{34} \\ 0 & 0 & 0 & u_{44} \end{bmatrix} \cdot \quad (21)$$

Przystępujemy do przeglądania drugiej kolumny:

$$a_{12} : \quad u_{12} = 4 \quad (22a)$$

$$a_{22} : \quad \frac{1}{2} \cdot 4 + u_{22} = 2 \implies u_{22} = 0 \quad (22b)$$

$$a_{32} : \quad 0 \cdot 4 + l_{32}u_{22} = 1 \implies l_{32}u_{22} = 1 \quad (22c)$$

$$a_{42} : \quad -\frac{1}{2} \cdot 4 + l_{42}u_{22} = 1 \implies l_{42}u_{22} = 3 \quad (22d)$$

Widać, iż równań (22) nie da się rozwiązać ze względu na l_{32} , l_{42} . Dzieje się tak dlatego, że aktualny element diagonalny („element podstawowy”) jest zerem. Aby uniknąć tej sytuacji, należy przestawić drugi wiersz faktoryzowanej macierzy z pewnym innym wierszem leżącym *poniżej* drugiego; oczywiście należy także przestawić już obliczone elementy macierzy \mathbf{L} odpowiadające przestawianym wierszom \mathbf{A} . Jako wiersz, który zajmie miejsce wiersza drugiego, wybieramy ten, który prowadzi do największej (na moduł) wartości po prawej stronie równań (22), jako że ta wartość stanie

się nowym elementem diagonalnym, przez który będziemy dzielić. W naszym przykładzie jest to wiersz czwarty. Zatem

$$\begin{bmatrix} 2 & 4 & 1 & 1 \\ -1 & 1 & 0 & 1 \\ 0 & 1 & 2 & -1 \\ 1 & 2 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 & 0 \\ 0 & l_{32} & 1 & 0 \\ \frac{1}{2} & l_{42} & l_{43} & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & u_{13} & u_{14} \\ 0 & u_{22} & u_{23} & u_{24} \\ 0 & 0 & u_{33} & u_{34} \\ 0 & 0 & 0 & u_{44} \end{bmatrix}. \quad (23)$$

Kolory wskazują co z czym było przestawiane. Podkreślam, iż w macierzy \mathbf{L} przestawieniu podlegają tylko *już obliczone* elementy, a więc elementy leżące na lewo od aktualnie analizowanej kolumny. Ponieważ wiersze leżące powyżej aktualnie obliczanego elementu diagonalnego nie ulegają zmianie, obliczoną wartość u_{12} można już było wpisać do macierzy. Teraz

z łatwością obliczamy

$$a_{22} : -\frac{1}{2} \cdot 4 + u_{22} = 1 \implies u_{22} = 3 \quad (24a)$$

$$a_{32} : 0 \cdot 4 + l_{32}u_{22} = 1 \implies l_{32} = \frac{1}{3} \quad (24b)$$

$$a_{42} : \frac{1}{2} \cdot 4 + l_{42}u_{22} = 2 \implies l_{42} = 0 \quad (24c)$$

a zatem

$$\begin{bmatrix} 2 & 4 & 1 & 1 \\ -1 & 1 & 0 & 1 \\ 0 & 1 & 2 & -1 \\ 1 & 2 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 & 0 \\ 0 & \frac{1}{3} & 1 & 0 \\ \frac{1}{2} & 0 & l_{43} & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & u_{13} & u_{14} \\ 0 & 3 & u_{23} & u_{24} \\ 0 & 0 & u_{33} & u_{34} \\ 0 & 0 & 0 & u_{44} \end{bmatrix}. \quad (25)$$

Przystępujemy do przeglądania trzeciej kolumny.

$$a_{13} : \quad u_{13} = 1 \quad (26a)$$

$$a_{23} : \quad -\frac{1}{2} \cdot u_{13} + u_{23} = 0 \implies u_{23} = \frac{1}{2} \quad (26b)$$

$$a_{33} : \quad 0 \cdot u_{13} + \frac{1}{3} \cdot u_{23} + u_{33} = 2 \implies u_{33} = \frac{11}{6} \quad (26c)$$

$$a_{43} : \quad \frac{1}{2} \cdot u_{13} + 0 \cdot u_{23} + l_{43}u_{33} = 3 \implies l_{43}u_{33} = \frac{5}{2} \quad (26d)$$

W tym wypadku nie *musimy* permutować wierszy (równania (26) nie zawierają dzielenia przez zero), tym niemniej *powinniśmy* to zrobić, aby elementem diagonalnym był element o możliwie największym module. Ponieważ $5/2 > 11/6$, permutujemy trzeci i czwarty wiersz macierzy \mathbf{A} , przedstawiając jednocześnie *już obliczone* elementy macierzy \mathbf{L} .

A zatem

$$\begin{bmatrix} 2 & 4 & 1 & 1 \\ -1 & 1 & 0 & 1 \\ 1 & 2 & 3 & 1 \\ 0 & 1 & 2 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 & 0 \\ \frac{1}{2} & 0 & 1 & 0 \\ 0 & \frac{1}{3} & l_{43} & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & 1 & u_{14} \\ 0 & 3 & \frac{1}{2} & u_{24} \\ 0 & 0 & u_{33} & u_{34} \\ 0 & 0 & 0 & u_{44} \end{bmatrix}. \quad (27)$$

Jak poprzednio, kolory pokazują elementy, które zostały przestawione. Teraz z łatwością obliczamy najpierw brakujące elementy u_{33} , l_{43} , później zaś elementy ostatniej kolumny macierzy U — w tym przypadku nie trzeba (a nawet nie da się) wykonywać już żadnych „pivotów”. Ostatecznie otrzymujemy

$$\begin{bmatrix} 2 & 4 & 1 & 1 \\ -1 & 1 & 0 & 1 \\ 1 & 2 & 3 & 1 \\ 0 & 1 & 2 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 & 0 \\ \frac{1}{2} & 0 & 1 & 0 \\ 0 & \frac{1}{3} & \frac{11}{15} & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & 1 & 1 \\ 0 & 3 & \frac{1}{2} & \frac{3}{2} \\ 0 & 0 & \frac{5}{2} & \frac{1}{2} \\ 0 & 0 & 0 & -\frac{28}{15} \end{bmatrix}. \quad (28)$$

Widać zatem, że

1. Faktoryzacja LU nie wymaga *de facto* rozwiązywania skomplikowanego układu równań, jako że każde z rozwiązywanych równań jest równaniem z jedną niewiadomą, jeśli tylko macierz A jest przeglądana we właściwej kolejności. Obliczenie jednego elementu wymaga $\sim N$ operacji, wszystkich elementów jest N^2 , zatem koszt obliczeniowy faktoryzacji LU jest rzędu $O(N^3)$.
2. Macierz A można przeglądać kolumnami poczynając od lewego górnego rogu, lecz jeszcze bardziej naturalna jest następująca kolejność:
 - (a) Zaczynamy od lewego górnego rogu.

- (b) Przeglądając k -tą kolumnę od pozycji diagonalnej w dół obliczamy wszystkie iloczyny $l_{kk}u_{kk}, l_{k+1,k}u_{kk}, \dots, l_{Nk}u_{kk}$ *bez wykonywania dzielenia przez u_{kk}* . Jako element podstawowy wybieramy ten z nich, który ma największą (na moduł) wartość — w tym celu przedstawiamy odpowiednie wiersze A oraz odpowiednie elementy L stojące w już obliczonych kolumnach $(1, \dots, k-1)$. Teraz wykonujemy dzielenie przez nowe u_{kk} ($l_{kk} = 1$). Widać, że iloczynów $l_{sk}u_{kk}, s > k$, nie trzeba ponownie obliczać, ponieważ zostały policzone przed wybraniem elementu podstawowego.
- (c) Po przejrzaniu k -tej kolumny przeglądamy k -ty wiersz poczynając od pozycji $k+1$ (poprzednie elementy tego wiersza zostały już obliczone przy okazji przeglądania poprzednich kolumn), jako że nie biorą one udziału w wyborze elementu podstawowego, wszystkie zaś elementy potrzebne do ich obliczenia są już w tym momencie znane.

3. Na skutek zastosowania wyboru elementu podstawowego dostajemy **nie** faktoryzację wyjściowej macierzy A , lecz faktoryzację macierzy różniące się od macierzy wyjściowej kolejnością wierszy (porównaj lewe strony (19) i (28)). Trzeba zapamiętać tę permutację wierszy, jako że przy rozwiązywaniu równania $Ax = b$ trzeba zastosować tę samą permutację elementów wektora b .

Algorytm Thomasa

Można łatwo zauważyć, że faktoryzacji LU macierzy trójdzielnej można dokonać w czasie liniowym. Istotnie, gdy obliczamy elementy $l_{ij}, i > j$, widzimy, że dla takiej macierzy tylko $l_{n,n+1} \neq 0$ — pozostałych elementów nie trzeba więc obliczać, skoro z góry wiadomo, że znikają. Czynniki L jest dwudzielny, podobnie dwudzielny jest czynnik U , a zatem także forward substitution i backsubstitution można wykonać w czasie liniowym. Uwaga: Ze względu na konieczność zachowania kształtu macierzy trójdzielnej, niemożliwy jest przy tym wybór elementu podstawowego.

Złożoność obliczeniowa rozwiązywania układu równań z macierzą trójdzielną wynosi $O(N)$.

Algorytm faktoryzacji LU macierzy trójdzielnej, wraz z forward substitution i backsubstitution, nosi nazwę algorytmu Thomasa.

Faktoryzacja Cholesky'ego

Niech $\mathbf{A} \in \mathbb{R}^{N \times N}$ będzie symetryczna, $\mathbf{A}^T = \mathbf{A}$, i dodatnio określona:

$$\forall \mathbf{x} \in \mathbb{R}^N, \mathbf{x} \neq \mathbf{0}: \mathbf{x}^T \mathbf{A} \mathbf{x} > 0. \quad (29)$$

Wówczas istnieje alternatywa dla faktoryzacji LU : faktoryzacja postaci

$$\mathbf{A} = \mathbf{C} \mathbf{C}^T, \quad (30)$$

gdzie \mathbf{C} jest macierzą trójkątną dolną o elementach diagonalnych większych od zera. Znalezienie faktoryzacji Cholesky'ego jest mniej więcej o połowę szybsze, niż znalezienie faktoryzacji LU tej samej macierzy.

Najprostszy algorytm jest bardzo podobny do algorytmu Doolittle'a:

$$\underbrace{\begin{bmatrix} c_{11} & & & & \\ c_{21} & c_{22} & & & \\ c_{31} & c_{32} & c_{33} & & \\ c_{41} & c_{42} & c_{43} & c_{44} & \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}}_C \underbrace{\begin{bmatrix} c_{11} & c_{21} & c_{31} & c_{41} & \dots \\ & c_{22} & c_{32} & c_{42} & \dots \\ & & c_{33} & c_{43} & \dots \\ & & & c_{44} & \dots \\ & & & & \ddots \end{bmatrix}}_{C^T} = \underbrace{\begin{bmatrix} a_{11} & a_{21} & a_{31} & a_{41} & \dots \\ a_{21} & a_{22} & a_{32} & a_{42} & \dots \\ a_{31} & a_{32} & a_{33} & a_{43} & \dots \\ a_{41} & a_{42} & a_{43} & a_{44} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}}_A \quad (31)$$

Pierwsza kolumna macierzy A daje

$$\begin{aligned}
 c_{11}^2 &= a_{11} \\
 c_{21}c_{11} &= a_{21} \\
 c_{31}c_{11} &= a_{31} \\
 c_{41}c_{11} &= a_{41} \\
 &\dots\dots\dots
 \end{aligned} \quad (32)$$

Z pierwszego z tych równań obliczamy c_{11} , z kolejnych c_{21}, c_{31} itd.

Druga kolumna daje

$$\begin{aligned}c_{11}c_{21} &= a_{21} \\c_{21}^2 + c_{22}^2 &= a_{22} \\c_{31}c_{21} + c_{32}c_{22} &= a_{32} \\c_{41}c_{21} + c_{42}c_{22} &= a_{42} \\&\dots\dots\dots\end{aligned}\tag{33}$$

Pierwsze z równań (33) jest identyczne z drugim z równań (32). Drugie z równań (33) pozwala na wyliczenie c_{22} . Dalsze równania pozwalają wyliczyć c_{32} , c_{42} itd.

I tak dalej.

Z uwagi na symetrię problemu, przy obliczaniu faktoryzacji Cholesky'ego nie jest możliwy wybór elementów podstawowych. Z uwagi na kolejność obliczeń, obliczone czynniki Cholesky'ego można przechowywać w tym samym miejscu, co elementy pierwotnej macierzy A .

Faktoryzacja LDL

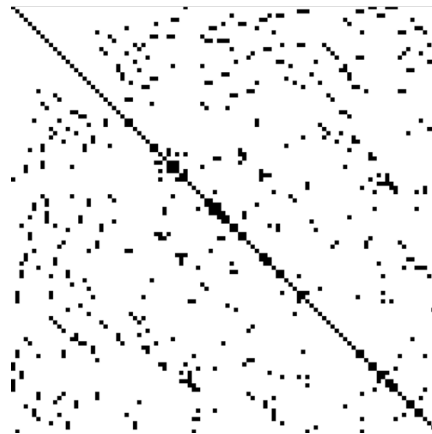
Jeżeli macierz spełnia założenia potrzebne do przeprowadzenia faktoryzacji Cholesky'ego, można także znaleźć jej inną faktoryzację:

$$\mathbf{A} = \mathbf{L} \mathbf{D} \mathbf{L}^T, \quad (34)$$

gdzie \mathbf{L} jest macierzą trójkątną dolną o tej własności, że $\forall i: l_{ii} = 1$, natomiast \mathbf{D} jest macierzą diagonalną o dodatnich elementach. Zaletą faktoryzacji LDL w stosunku do faktoryzacji Cholesky'ego jest to, iż do znalezienia LDL nie potrzeba pierwiastkowań.

Macierze rzadkie

W wielu praktycznych zastosowaniach występują **macierze rzadkie**, to znaczy takie, w których liczba elementów niezerowych rośnie wolniej niż N^2 , gdzie N jest wymiarem macierzy. Na przykład w macierzy trójdzielnej liczba niezerowych elementów skaluje się jak $O(3N)$, a w macierzy pasmowej o P dodatkowych diagonalach jak $O((2P + 1)N)$. Możliwe są także inne struktury macierzy rzadkich.



Macierze rzadkie i efektywność numeryczna

Dla efektywności numerycznej jest niesłuchanie ważne, aby zastosowany **algorytm** uwzględniał strukturę macierzy, tak, aby nie trzeba było wykonywać redundantnych mnożeń przez zero i dodawań zera, a nawet żeby nie przechodzić przez zerowe elementy.

- Dla macierzy trójdzielnej faktoryzacji LU dokonujemy w czasie liniowym, $O(N)$, ale za to niemożliwy jest wybór elementu podstawowego.
- Jeżeli możliwa jest faktoryzacja Cholesky'ego macierzy M -diagonalnej, także jej czynnik Cholesky'ego będzie M -diagonalny. Może jednak pojawić się niekorzystne zjawisko, zwane **wypełnieniem**: Jeżeli sama macierz ma zera “wewnątrz” pasma, jej czynnik Cholesky'ego nie musi ich mieć, co może bardzo niekorzystnie wpłynąć na wydajność numeryczną.

Przykład

Czynnik Cholesky'ego następującej macierzy *rzadkiej*

$$\begin{bmatrix} \bullet & \bullet & \bullet & \bullet & \bullet & \dots \\ \bullet & \bullet & & & & \\ \bullet & & \bullet & & & \\ \bullet & & & \bullet & & \\ \bullet & & & & \bullet & \\ \vdots & & & & & \ddots \end{bmatrix}$$

(35)

(niewypełnione elementy są zerami) będzie macierzą *pełną*.

Faktoryzacja QR

Innym rodzajem faktoryzacji jest faktoryzacja QR : daną macierz $A \in \mathbb{R}^{N \times N}$ przedstawiam w postaci iloczynu

$$A = QR, \quad (36)$$

gdzie Q jest macierzą ortogonalną, a R jest macierzą trójkątną górną. Złożoność obliczeniowa faktoryzacji QR wynosi dla macierzy pełnych $O(N^3)$, czyli tyle samo, co faktoryzacji LU , jednak współczynnik przy wyrazie wiodącym jest gorszy niż dla LU . QR nie jest więc metodą “z wyboru” rozwiązywania układów równań liniowych. Jeśli jednak z jakichś względów

faktoryzację QR możemy łatwo (lub musimy) obliczyć, układ równań liniowych rozwiązujemy jak następuje:

$$\mathbf{Ax} = \mathbf{b} \quad (37a)$$

$$\mathbf{QRx} = \mathbf{b} \quad (37b)$$

$$\mathbf{Rx} = \mathbf{Q}^T \mathbf{b} \quad (37c)$$

Koszt obliczeniowy przejścia od (37b) do (37c) wynosi $O(N^2)$. Równanie (37c) rozwiązujemy metodą *backsubstitution*, co także kosztuje $O(N^2)$. Jest to więc koszt mały w porównaniu z dokonaniem samej faktoryzacji.

Pozostaje pytanie: Jak dokonać tej faktoryzacji?

Transformacja Householdera

Niech $\mathbf{u} \in \mathbb{R}^N$, $\mathbf{u} \neq 0$. Tworzymy macierz

$$\mathbf{P} = \mathbb{I} - 2 \frac{\mathbf{u}\mathbf{u}^T}{\|\mathbf{u}\|^2}. \quad (38)$$

W sposób oczywisty $\mathbf{P}^T = \mathbf{P}$. Obliczmy

$$\begin{aligned} \mathbf{P}^2 &= \left(\mathbb{I} - 2 \frac{\mathbf{u}\mathbf{u}^T}{\|\mathbf{u}\|^2} \right) \left(\mathbb{I} - 2 \frac{\mathbf{u}\mathbf{u}^T}{\|\mathbf{u}\|^2} \right) \\ &= \mathbb{I} - 2 \frac{\mathbf{u}\mathbf{u}^T}{\|\mathbf{u}\|^2} - 2 \frac{\mathbf{u}\mathbf{u}^T}{\|\mathbf{u}\|^2} + 4 \frac{\overbrace{\mathbf{u}\mathbf{u}^T\mathbf{u}\mathbf{u}^T}^{\|\mathbf{u}\|^2}}{\|\mathbf{u}\|^4} \\ &= \mathbb{I} - 4 \frac{\mathbf{u}\mathbf{u}^T}{\|\mathbf{u}\|^2} + 4 \frac{\mathbf{u}\mathbf{u}^T}{\|\mathbf{u}\|^2} \\ &= \mathbb{I} \end{aligned} \quad (39)$$

Skoro $\mathbf{P}^T = \mathbf{P}$ oraz $\mathbf{P}^2 = \mathbb{I}$, $\mathbf{P}^T = \mathbf{P}^{-1}$: macierz (38) jest macierzą symetryczną, rzeczywistą oraz ortogonalną. Macierz taką nazywamy ortogonalną macierzą rzutową.

Niech teraz w (38)

$$\mathbf{u} = \mathbf{x} \mp \|\mathbf{x}\| \hat{\mathbf{e}}_1, \quad (40)$$

gdzie $\hat{\mathbf{e}}_1$ jest pierwszym wektorem jednostkowym. Macierz (38) wraz z (40) nazywam *macierzą Householdera*. Obliczam

$$\mathbf{P}\mathbf{x} = \mathbf{x} - \frac{2\mathbf{u}\mathbf{u}^T\mathbf{x}}{\|\mathbf{u}\|^2}. \quad (41a)$$

Zauważmy, że $\mathbf{u}^T\mathbf{x} = \mathbf{x}^T\mathbf{x} \mp \|\mathbf{x}\| \hat{\mathbf{e}}_1^T\mathbf{x} = \|\mathbf{x}\|^2 \mp \|\mathbf{x}\| x_1$, gdzie x_1 jest pierwszą składową wektora \mathbf{x} .

Analogicznie $\|\mathbf{u}\|^2 = (\mathbf{x} \mp \|\mathbf{x}\| \hat{\mathbf{e}}_1)^T (\mathbf{x} \mp \|\mathbf{x}\| \hat{\mathbf{e}}_1) = \mathbf{x}^T \mathbf{x} \mp \|\mathbf{x}\| x_1 \mp \|\mathbf{x}\| x_1 + \|\mathbf{x}\|^2 = 2 (\|\mathbf{x}\|^2 \mp \|\mathbf{x}\| x_1)$. Wobec tego

$$\mathbf{P}_\mathbf{x} = \mathbf{x} - \frac{2\mathbf{u} (\|\mathbf{x}\|^2 \mp \|\mathbf{x}\| x_1)}{2 (\|\mathbf{x}\|^2 \mp \|\mathbf{x}\| x_1)} = \mathbf{x} - \mathbf{u} = \pm \|\mathbf{x}\| \hat{\mathbf{e}}_1. \quad (41b)$$

Efektem działania macierzy Householdera na wskazany wektor jest wyzerowanie wszystkich jego składowych, poza pierwszą, i “przelanie” całej jego długości na pierwszą składową. Złożoność obliczeniowa tej procedury wynosi $O(N)$.

Faktoryzacja QR

Niech $\mathbf{A} \in \mathbb{R}^{N \times N}$. Niech \mathbf{P}_1 oznacza transformację Householdera zbudowaną na pierwszej kolumnie macierzy \mathbf{A} . Otrzymujemy

$$\mathbf{P}_1 \mathbf{A} = \mathbf{A}_1 \quad (42)$$
$$\mathbf{P}_1 \begin{bmatrix} \bullet & \bullet & \bullet & \bullet & \cdots \\ \bullet & \bullet & \bullet & \bullet & \cdots \\ \bullet & \bullet & \bullet & \bullet & \cdots \\ \bullet & \bullet & \bullet & \bullet & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} = \begin{bmatrix} \bullet & \bullet & \bullet & \bullet & \cdots \\ & \bullet & \bullet & \bullet & \cdots \\ & \bullet & \bullet & \bullet & \cdots \\ & \bullet & \bullet & \bullet & \cdots \\ & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Transformacja Householdera \mathbf{P}_1 wyzerowała pierwszą kolumnę macierzy \mathbf{A} , za wyjątkiem elementu diagonalnego. Złożoność obliczeniowa tego kroku wynosi $O(N^2)$ (transformacja Householdera działa na N kolumn).

Niech teraz

$$\mathbf{P}_2 = \left[\begin{array}{c|cccc} 1 & 0 & 0 & \dots & 0 \\ \hline 0 & & & & \\ 0 & & & & \\ 0 & & & & \\ 0 & & & & \end{array} \right] \quad (43)$$

$(N-1)\mathbf{P}_2$

gdzie $(N-1)\mathbf{P}_2 \in \mathbb{R}^{(N-1) \times (N-1)}$ jest transformacją Householdera, zbudowaną na drugiej kolumnie macierzy \mathbf{A}_1 , poczynając od elementu diagonalnego w dół. Otrzymujemy

$$\mathbf{P}_2\mathbf{P}_1\mathbf{A} = \mathbf{P}_2\mathbf{A}_1 = \mathbf{A}_2 = \begin{bmatrix} \bullet & \bullet & \bullet & \bullet & \dots \\ & \bullet & \bullet & \bullet & \dots \\ & & \bullet & \bullet & \dots \\ & & \bullet & \bullet & \dots \\ & & \vdots & \vdots & \ddots \end{bmatrix} \quad (44)$$

Następnie definiuję

$$\mathbf{P}_3 = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 \\ \hline 0 & 0 & & & & \\ 0 & 0 & & & & \\ 0 & 0 & & & & \\ 0 & 0 & & & & \end{array} \right] \quad (45)$$

gdzie $(N-2)\mathbf{P}_3 \in \mathbb{R}^{(N-2) \times (N-2)}$ jest transformacją Householdera, zbudowaną na trzeciej kolumnie macierzy \mathbf{A}_2 , poczynając od elementu diagonalnego w dół. Stojąca w lewym górnym rogu macierz jednostkowa służy do tego, żeby nie zepsuć struktury, którą osiągnęliśmy w poprzednich kro-

kach. Otrzymujemy

$$\mathbf{P}_3\mathbf{P}_2\mathbf{P}_1\mathbf{A} = \mathbf{P}_3\mathbf{A}_2 = \mathbf{A}_3 = \begin{bmatrix} \bullet & \bullet & \bullet & \bullet & \cdots \\ & \bullet & \bullet & \bullet & \cdots \\ & & \bullet & \bullet & \cdots \\ & & & \bullet & \cdots \\ & & & \vdots & \ddots \end{bmatrix} \quad (46)$$

Widać, że po $N-1$ krokach osiągnę

$$\mathbf{R} = \begin{bmatrix} \bullet & \bullet & \bullet & \cdots & \bullet \\ & \bullet & \bullet & \cdots & \bullet \\ & & \bullet & \cdots & \bullet \\ & & & \ddots & \vdots \\ & & & & \bullet \end{bmatrix} = \underbrace{\mathbf{P}_{N-1}\mathbf{P}_{N-2}\cdots\mathbf{P}_1}_{\mathbf{Q}^T} \mathbf{A} \quad (47)$$

\mathbf{R} jest macierzą trójkątną górną. Ponieważ macierze \mathbf{P}_i są ortogonalne, ich iloczyn, oznaczony przez \mathbf{Q}^T , także jest macierzą ortogonalną. Nie

musimy zapamiętywać poszczególnych macierzy P_i , wystarczy zapamiętać ich iloczyn.

Otrzymaliśmy zatem dla dowolnej macierzy kwadratowej *faktoryzację na macierz ortogonalną i trójkątną górną*:

$$A = QR, \quad (48)$$

czyli poszukiwaną faktoryzację QR .

Nie jest to jedyny algorytm pozwalający uzyskać faktoryzację QR , pokazuje jednak, że jest ona w ogólnym (czyli dowolnym) wypadku możliwa.

Obroty Givensa

Transformacja Householdera służy do zerowania wielu składowych jakiegoś wektora. Jeżeli chcemy selektywnie wyzerować jakieś składowe — lub jeśli interesujący nas wektor ma jakąś szczególną postać — bardziej efektywne od transformacji Householdera będą *obroty Givensa*.

Niech x będzie pewnym wektorem i niech $y = G(i, j)x$. Składowe wektora y wynoszą

$$y_k = \begin{cases} cx_i + sx_j & k = i \\ -sx_i + cx_j & k = j \\ x_k & \text{poza tym} \end{cases} \quad (50)$$

Zażądajmy, aby $y_j = 0$. Widać, że musi zachodzić

$$c = \frac{x_i}{\sqrt{x_i^2 + x_j^2}}, \quad s = \frac{x_j}{\sqrt{x_i^2 + x_j^2}}. \quad (51)$$

Obrót Givensa (49) wraz z warunkami (51) zeruje j -tą składową wybranego wektora. Składowa i -ta przybiera wartość $\sqrt{x_i^2 + x_j^2}$.

Faktoryzacja QR macierzy trójdzielnej symetrycznej

Rozpatrzmy macierz $\mathbf{A} \in \mathbb{R}^{N \times N}$, trójdzielną symetryczną

$$\mathbf{A} = \begin{bmatrix} a & b & & & & \\ b & d & e & & & \\ & e & f & g & & \\ & & g & h & l & \\ & & & \ddots & \ddots & \ddots \end{bmatrix} \quad (52)$$

Zadziałajmy na nią macierzą Givensa taką, aby zerowała drugi element pierwszej kolumny

$$\mathbf{G}_1 = \begin{bmatrix} c_1 & s_1 & & & & \\ -s_1 & c_1 & & & & \\ & & 1 & & & \\ & & & 1 & & \\ & & & & \ddots & \end{bmatrix} \quad (53)$$

$$\mathbf{A}_1 = \mathbf{G}_1 \mathbf{A} = \begin{bmatrix} \sqrt{a^2 + b^2} & \frac{ab+bd}{\sqrt{a^2+b^2}} & \frac{be}{\sqrt{a^2+b^2}} & & & & & & \\ & \frac{ad-b^2}{\sqrt{a^2+b^2}} & \frac{ae}{\sqrt{a^2+b^2}} & & & & & & \\ & e & f & g & & & & & \\ & & g & h & l & & & & \\ & & & \dots & \dots & \dots & & & \end{bmatrix} \quad (54)$$

Wiersze macierzy \mathbf{A}_1 począwszy od trzeciego w dół zgadzają się z wierszami macierzy \mathbf{A} . Obliczenie macierzy \mathbf{A}_1 wymaga wykonania stałej, *niezależnej od rozmiaru macierzy*, liczby operacji. W wyniku otrzymaliśmy macierz, w której poddiagonalne elementy pierwszej kolumny są zerami.

Macierz A_1 mnożymy przez macierz Givensa

$$G_2 = \begin{bmatrix} 1 & & & & & \\ & c_2 & s_2 & & & \\ & -s_2 & c_2 & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & \ddots \end{bmatrix} \quad (55)$$

dobraną tak, aby zerowała trzeci element drugiej kolumny macierzy A_1 . Pierwszy wiersz i pierwsza kolumna nie zmieniają się, podobnie jak wiersze począwszy od czwartego. W rezultacie macierz $A_2 = G_2 A_1 = G_2 G_1 A$ ma zera w poddiagonalnych miejscach dwu pierwszych kolumn. Ten krok także wymaga stałej, niezależnej od rozmiaru macierzy, liczby operacji.

W kolejnym kroku macierz A_2 mnożymy przez taką macierz Givensa, która wyzeruje czwarty element trzeciej kolumny. I tak dalej.

W ten sposób, po $N-1$ krokach, ponosząc koszt numeryczny $O(N)$ (stały koszt na krok, $\sim N$ kroków), otrzymujemy

$$\mathbf{G}_{N-1} \cdots \mathbf{G}_2 \mathbf{G}_1 \mathbf{A} = \mathbf{R} = \begin{bmatrix} \bullet & \bullet & \bullet & & & & \\ & \bullet & \bullet & \bullet & & & \\ & & \bullet & \bullet & \bullet & & \\ & & & \cdots & \cdots & \cdots & \\ & & & & \bullet & \bullet & \\ & & & & & \bullet & \bullet \\ & & & & & & \bullet \end{bmatrix}, \quad (56a)$$

czyli

$$\mathbf{A} = \underbrace{\mathbf{G}_1^T \mathbf{G}_2^T \cdots \mathbf{G}_{N-1}^T}_{\mathbf{Q}} \mathbf{R} \quad (56b)$$

Macierz \mathbf{Q} jest ortogonalna. Macierz \mathbf{R} jest trójkątna górna (tak naprawdę ma ona tylko dwie niezerowe diagonale nad diagonalą główną). Widzimy, że (56b) jest faktoryzacją QR macierzy trójdzielnej symetrycznej.

Zastosowanie do rozwiązywania układu równan liniowych

Jeżeli chcemy użyć obrotów Givensa do rozwiązania układu równań liniowych

$$\mathbf{Ax} = \mathbf{b}, \quad (57)$$

gdzie \mathbf{A} jest trójdagonalną macierzą symetryczną, postępując jak poprzednio otrzymujemy kolejno

$$\mathbf{G}_1 \mathbf{Ax} = \mathbf{G}_1 \mathbf{b} \quad (58a)$$

$$\mathbf{G}_2 \mathbf{G}_1 \mathbf{Ax} = \mathbf{G}_2 \mathbf{G}_1 \mathbf{b} \quad (58b)$$

...

$$\mathbf{G}_{N-1} \cdots \mathbf{G}_2 \mathbf{G}_1 \mathbf{Ax} \equiv \mathbf{Rx} = \mathbf{G}_{N-1} \cdots \mathbf{G}_2 \mathbf{G}_1 \mathbf{b}. \quad (58c)$$

Oczywiście istotne jest tylko równanie (58c) — lewych stron poprzednich równań nie musimy wyliczać. Każde kolejne mnożenie po stronie prawej

wykonujemy w stałym czasie, a więc do postaci $\mathbf{R}\mathbf{x} = \tilde{\mathbf{b}}$ dochodzimy w czasie $O(N)$. To równanie rozwiązujemy metodą *backsubstitution*, co, z uwagi na szczególną postać macierzy \mathbf{R} , także da się wykonać w czasie liniowym.

Przykład ten pokazuje, że możemy odnieść duży zysk na złożoności obliczeniowej, jeśli tylko dobierzemy odpowiedni algorytm odpowiadający strukturze — w tym wypadku rzadkości i symetryczności — macierzy.

Uwaga: Skumulowanej macierzy Givensa \mathbf{Q} nie musimy wyliczać w sposób jawny — gdybyśmy to chcieli zrobić, wymagałoby to $O(N^2)$ operacji.

Równania macierzowe

Przypuśćmy, że mamy rozwiązać kilka układów równań z tą samą lewą stroną, a różnymi wyrazami wolnymi:

$$\mathbf{A}\mathbf{x}^{(i)} = \mathbf{b}^{(i)}, \quad i = 1, 2, \dots, M \quad (59)$$

gdzie $\mathbf{A} \in \mathbb{R}^{N \times N}$, $\mathbf{x}^{(i)}, \mathbf{b}^{(i)} \in \mathbb{R}^N$. Zauważmy, że korzystając z własności mnożenia macierzy, równania (59) można zapisać w postaci

$$\mathbf{A}\mathbf{X} = \mathbf{B} \quad (60)$$

gdzie $\mathbf{A} \in \mathbb{R}^{N \times N}$, $\mathbf{X}, \mathbf{B} \in \mathbb{R}^{N \times M}$, przy czym

$$\mathbf{X} = \begin{bmatrix} x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(M)} \\ x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(M)} \\ \dots & \dots & \dots & \dots \\ x_N^{(1)} & x_N^{(2)} & \dots & x_N^{(M)} \end{bmatrix}$$

i analogicznie dla \mathbf{B} . Innymi słowy, **macierzowy układ równań (60) jest równoważny układowi równań liniowych (59) z M niezależnymi prawymi stronami**. Oczywiście do rozwiązywania układów równań macierzowych postaci (60) można używać wszystkich poznanych dotąd faktoryzacji, a także eliminacji Gaussa, gdyż znamy z góry wszystkie prawe strony.

Uwaga — jawna konstrukcja macierzy odwrotnej

Z powyższych uwag widać, że problem jawnej konstrukcji macierzy odwrotnej

$$\mathbf{A} \cdot \mathbf{A}^{-1} = \mathbb{I} \quad (61)$$

jest problemem postaci (60), a więc kolejne kolumny macierzy odwrotnej uzyskujemy rozwiązując kolejne układy (59) dla $i = 1, 2, \dots, N$, przy czym $\mathbf{b}^{(1)} = [1, 0, 0, \dots]^T$, $\mathbf{b}^{(2)} = [0, 1, 0, \dots]^T$ itd. Rozwiązanie układu równań $\mathbf{A}\mathbf{x} = \mathbf{b}$ poprzez jawną konstrukcję macierzy odwrotnej, $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$, **wymaga rozwiązania N układów równań liniowych, co oznacza koszt $O(2N^3)$** , podczas gdy koszt bezpośredniego rozwiązania równania $\mathbf{A}\mathbf{x} = \mathbf{b}$ to $O(N^3)$.

Pojawiający się często we wzorach napis $A^{-1}b$ **zawsze** rozumiemy jako wezwanie do znalezienie wektora z takiego, że $Az = b$, **nigdy** nie jako polecenie jawnego skonstruowania macierzy A^{-1} .

Przykład

Wyrażenie

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \mathbf{J}^{-1}\mathbf{f}_n \quad (62)$$

interpretujemy jako

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \mathbf{z} \quad (63a)$$

gdzie

$$\mathbf{J}\mathbf{z} = \mathbf{f}_n \quad (63b)$$

Wzór Shermana-Morrisona

Twierdzenie: Niech $\mathbf{A} \in \mathbb{R}^{N \times N}$, $\det \mathbf{A} \neq 0$ oraz $\mathbf{u}, \mathbf{v} \in \mathbb{R}^N$. Niech $\mathbf{A}_1 = \mathbf{A} + \mathbf{u}\mathbf{v}^T$. Wówczas

$$\mathbf{A}_1^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}}{1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}}. \quad (64)$$

Zauważmy, że ponieważ $\det \mathbf{A} \neq 0$, macierz \mathbf{A}^{-1} istnieje. Ponadto wyrażenie $\mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}$ jest *liczbą* (skalarem).

Wzór ten jest przydatny w przypadku, gdy chcemy wyliczyć \mathbf{A}_1^{-1} , gdzie \mathbf{A}_1 jest pewną szczególną modyfikacją macierzy \mathbf{A} , a odwrotność \mathbf{A}^{-1} znamy.

Przykład

Niech $\mathbf{u} = \mathbf{v} = [1, 0, 0, 0, 1]^T$. Wówczas

$$\mathbf{u}\mathbf{v}^T = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} [1 \ 0 \ 0 \ 0 \ 1] = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (65)$$

Niech teraz

$$\mathbf{A} = \begin{bmatrix} 3 & 1 & 0 & 0 & 0 \\ 1 & 4 & 1 & 0 & 0 \\ 0 & 1 & 4 & 1 & 0 \\ 0 & 0 & 1 & 4 & 1 \\ 0 & 0 & 0 & 1 & 3 \end{bmatrix}, \quad \mathbf{A}_1 = \mathbf{A} + \mathbf{u}\mathbf{v}^T = \begin{bmatrix} 4 & 1 & 0 & 0 & 1 \\ 1 & 4 & 1 & 0 & 0 \\ 0 & 1 & 4 & 1 & 0 \\ 0 & 0 & 1 & 4 & 1 \\ 1 & 0 & 0 & 1 & 4 \end{bmatrix}. \quad (66)$$

Dowód.

$$\begin{aligned}
 & (\mathbf{A} + \mathbf{u}\mathbf{v}^T) \left(\mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}}{1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}} \right) \\
 = & \mathbf{A}\mathbf{A}^{-1} - \frac{1}{1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}} \mathbf{A}\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1} + \mathbf{u}\mathbf{v}^T\mathbf{A}^{-1} \\
 & - \frac{1}{1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}} \mathbf{u} \underbrace{\mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}}_{\text{to jest liczba!}} \mathbf{v}^T\mathbf{A}^{-1} \\
 = & \mathbb{I} - \frac{1}{1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}} \mathbf{u}\mathbf{v}^T\mathbf{A}^{-1} + \mathbf{u}\mathbf{v}^T\mathbf{A}^{-1} - \frac{\mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}}{1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}} \mathbf{u}\mathbf{v}^T\mathbf{A}^{-1} \\
 = & \mathbb{I} + \left(1 - \frac{1}{1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}} - \frac{\mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}}{1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}} \right) \mathbf{u}\mathbf{v}^T\mathbf{A}^{-1} = \mathbb{I}. \quad (67)
 \end{aligned}$$

□

Algorytm Shermana-Morrisona

Wzór Shermana-Morrisona (64) pozwala skonstruować odwrotność macierzy A_1 jeśli znamy odwrotność A . Jednak w praktyce prawie **nigdy nie konstruujemy jawnej odwrotności macierzy!** Jak więc zastosować ten wzór?

Zauważmy, że **zapewne** chcemy obliczyć jakieś $A_1^{-1}b$, gdzie b jest znanym wektorem, przy założeniu, że **łatwo** potrafimy obliczyć $A^{-1}b$. Interesuje nas znalezienie

$$w = A_1^{-1}b = \left(A^{-1} - \frac{A^{-1}uv^T A^{-1}}{1 + v^T A^{-1}u} \right) b \quad (68)$$

Algorytm wygląda następująco:

(a) Rozwiąż równanie

$$\mathbf{Az} = \mathbf{b} \quad (69a)$$

(b) Rozwiąż równanie

$$\mathbf{Aq} = \mathbf{u} \quad (69b)$$

(c) Oblicz

$$\mathbf{w} = \mathbf{z} - \frac{\mathbf{v}^T \mathbf{z}}{1 + \mathbf{v}^T \mathbf{q}} \mathbf{q}. \quad (69c)$$

Problem sprowadza się więc do rozwiązania dwu równań (69a),(69b) z taką samą macierzą, które umiemy szybko rozwiązać, gdyż — na przykład — znamy faktoryzację macierzy \mathbf{A} . Zauważmy, że $\mathbf{v}^T \mathbf{A}^{-1} \mathbf{b} = \mathbf{v}^T \mathbf{z}$ jest *liczbą*.

Przykład (c.d.)

Pierwsza z macierzy (66) jest macierzą symetryczną, dodatnio określoną i trójdziagonalną, a więc jej czynnik Cholesky'ego ma tylko dwie niezerowe diagonale, a koszt jego wyliczenia jest rzędu $O(N)$. Czynnik Cholesky'ego drugiej z tych macierzy jest pełną macierzą trójkątną (nastąpi *wypełnienie*) i koszt jego wyliczenia jest rzędu $O(N^3)$ (wyobraźmy sobie, że zamiast o macierzach 5×5 , mówimy o macierzach 1000×1000). Zastosowanie algorytmu (69) redukuje problem do znalezienia i dwukrotnego zastosowania rzadkiego czynnika Cholesky'ego pierwszej z macierzy (66). Da się to zrobić w czasie liniowym.

Osobliwe układy równań

Dany jest układ równań

$$\begin{bmatrix} \frac{7}{6} & -\frac{1}{3} & -\frac{5}{6} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{5}{6} & -\frac{1}{3} & \frac{7}{6} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}. \quad (70)$$

Jak łatwo sprawdzić, wyznacznik główny tego układu równań wynosi zero, a więc układ równań (70) nie ma jednoznacznego rozwiązania. Ale może ma on rozwiązanie niejednoznaczne? Aby tak było, wszystkie jego wyznaczniki poboczne powinny zniknąć. Sprawdzamy:

$$\det \begin{bmatrix} 1 & -\frac{1}{3} & -\frac{5}{6} \\ 1 & \frac{2}{3} & -\frac{1}{3} \\ 1 & -\frac{1}{3} & \frac{7}{6} \end{bmatrix} = 2 \neq 0, \quad (71)$$

a więc układ równań (70) jest sprzeczny (nie ma rozwiązań). Z drugiej strony układ równań

$$\begin{bmatrix} \frac{7}{6} & -\frac{1}{3} & -\frac{5}{6} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{5}{6} & -\frac{1}{3} & \frac{7}{6} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \quad (72)$$

ma rozwiązanie, ale jest ono niejednoznaczne (wszystkie wyznaczniki pomocnicze równają się zero).

Układy równań (70) i (72) różnią się tylko prawymi stronami: ich macierz jest taka sama (i jest osobliwa). Widzimy, że w przypadku osobliwej macierzy układu równań to prawa strona decyduje, czy układ ma rozwiązania, czy też ich nie ma; *domyślamy się*, że to, jakie prawe strony oznaczają istnienie rozwiązań, jakie zaś ich brak, musi jakoś być związane z własnościami (osobliwej) macierzy układu równań. *Wygodnie byłoby mieć metodę pozwalającą na **zidentyfikowanie** wektorów, dla których osobliwy układ równań ma rozwiązania* (wzory Cramera pozwalają jedynie *sprawdzić*, czy dana prawa strona prowadzi do istnienia rozwiązań). *Metodą taką jest **SVD**.*

Singular Value Decomposition

Twierdzenie 1. *Dla każdej macierzy $\mathbf{A} \in \mathbb{R}^{M \times N}$, $M \geq N$, istnieje faktoryzacja*

$$\mathbf{A} = \mathbf{U} [\text{diag}(w_i)] \mathbf{V}^T, \quad (73)$$

*gdzie $\mathbf{U} \in \mathbb{R}^{M \times N}$ jest macierzą kolumnowo ortogonalną, $\mathbf{V} \in \mathbb{R}^{N \times N}$ jest macierzą ortogonalną oraz $w_i \in \mathbb{R}$, $i = 1, \dots, N$. Rozkład ten nazywamy rozkładem względem wartości osobliwych (*Singular Value Decomposition, SVD*). Jeżeli $M = N$, macierz \mathbf{U} jest macierzą ortogonalną.*

Nie przedstawiamy tu algorytmu tej faktoryzacji — w razie potrzeby można skorzystać z którejś ze sprawdzonych procedur bibliotecznych.

Złożoność obliczeniowa faktoryzacji *SVD* wynosi $O(N^3)$, ale współczynnik przy wyrazie wiodącym jest wysoki, więc nie jest to metoda “z wyboru” do rozwiązywania układów równań. Znajduje ona jednak ważne zastosowania w przypadku osobliwych lub źle uwarunkowanych układów równań. Aby to zrozumieć, trzeba najpierw omówić algebraiczne własności tej faktoryzacji.

Jądro i zasięg operatora

Niech $\mathbf{A} \in \mathbb{R}^{M \times N}$. *Jądrem operatora \mathbf{A}* nazywam

$$\text{Ker } \mathbf{A} = \{ \mathbf{x} \in \mathbb{R}^N : \mathbf{A}\mathbf{x} = \mathbf{0} \}. \quad (74)$$

Zasięgiem operatora \mathbf{A} nazywam

$$\text{Range } \mathbf{A} = \{ \mathbf{y} \in \mathbb{R}^M : \exists \mathbf{x} \in \mathbb{R}^N : \mathbf{A}\mathbf{x} = \mathbf{y} \}. \quad (75)$$

Jądro i zasięg operatora są przestrzeniami liniowymi. Jeśli $M = N < \infty$,
 $\dim(\text{Ker } \mathbf{A}) + \dim(\text{Range } \mathbf{A}) = N$.

Sens SVD

Sens *SVD* najlepiej widać w przypadku, w którym co najmniej jedna z wartości $w_i = 0$. Aby to zrozumieć, zobaczmy jak macierz $\mathbf{A} \in \mathbb{R}^{M \times N}$, posiadająca faktoryzację (73), działa na pewien wektor $\mathbf{x} \in \mathbb{R}^N$:

$$\mathbf{Ax} = \mathbf{U} [\text{diag}(w_i)] \mathbf{V}^T \mathbf{x} \quad (76)$$

Dla ustalenia uwagi, niech $w_1 = 0$, $w_{i>1} \neq 0$.

Przypomnienie z algebry ☹

Macierz $\mathbf{V} \in \mathbb{R}^{N \times N}$ jest macierzą *ortogonalną*, co oznacza, że jej kolejne kolumny stanowią ortonormalną bazę w \mathbb{R}^N . Wektor

$$\mathbf{z} = \mathbf{V}^T \mathbf{x} \quad (77)$$

stanowi rozkład wektora \mathbf{x} w bazie kolumn macierzy \mathbf{V} . Wyrażenie (76) możemy przepisać jako

$$\mathbf{Ax} = \mathbf{U} [\text{diag}(w_i)] \mathbf{z}, \quad (78)$$

gdzie \mathbf{z} spełnia (77).

Sens SVD — przypadek 1

Przypuśćmy, że wektor \mathbf{x} jest równy pierwszej kolumnie macierzy \mathbf{V} . Wówczas $\mathbf{z} = [1, 0, 0, \dots, 0]^T$, gdyż \mathbf{x} jest prostopadły do wszystkich kolumn \mathbf{V} poczynając od drugiej. W takim wypadku

$$\mathbf{Ax} = \mathbf{U} \begin{bmatrix} 0 & & & & & \\ & w_2 & & & & \\ & & w_3 & & & \\ & & & \ddots & & \\ & & & & w_N & \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (79)$$

Zero z lewego górnego rogu macierzy diagonalnej zabije jedynekę w wektorze po prawej stronie tej równości, a z kolei zera z tego wektora zabiją niezerowe w_i . Ostatecznie $\mathbf{Ax} = \vec{0}$. Stało się tak dlatego, że wektor \mathbf{x} był równy kolumnie macierzy \mathbf{V} , odpowiadającej zerowemu współczynnikowi w_i . *Kolumny macierzy \mathbf{V} , odpowiadające zerowym współczynnikom w_i , stanowią bazę w jądrze operatora \mathbf{A} .*

Sens SVD — przypadek 2

Niech teraz wektor \mathbf{x} będzie “dowolny”, to znaczy nie pozostaje w jakiejś specjalnej relacji do macierzy \mathbf{V} . Wówczas $\mathbf{z} = \mathbf{V}^T \mathbf{x} = [z_1, z_2, z_3, \dots, z_N]^T$, a zatem z (78) mamy

$$\mathbf{Ax} = \mathbf{U} \begin{bmatrix} 0 & & & & & \\ & w_2 & & & & \\ & & w_3 & & & \\ & & & \dots & & \\ & & & & w_N & \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \\ z_N \end{bmatrix} = \mathbf{U} \begin{bmatrix} 0 \\ w_2 z_2 \\ w_3 z_3 \\ \vdots \\ w_N z_N \end{bmatrix}. \quad (80)$$

Wynikiem ostatniego mnożenia będzie pewien wektor z przestrzeni \mathbb{R}^M . Ponieważ pierwszym elementem wektora $[0, w_2 z_2, \dots, w_N z_N]^T$ jest zero, **wynik ten nie zależy od pierwszej kolumny macierzy \mathbf{U}** . Kolumny macierzy \mathbf{U} są wzajemnie ortogonalne, a wektor \mathbf{Ax} nie zawiera przyczynku wzdłuż wektora będącego pierwszą kolumną \mathbf{U} , a jedynie przyczynki od wektorów

prostopadłych do tej kolumny. Widzimy zatem, że *kolumny macierzy U , odpowiadające niezerowym współczynnikom w_i , stanowią bazę w zasięgu operatora A .*

Dla macierzy osobliwych sens SVD sprowadza się do konstrukcji baz w jądrze i w zasięgu takiej macierzy. Jak za chwilę zobaczymy, można to uogólnić na przypadek macierzy nieosobliwych, ale bardzo źle uwarunkowanych.

Przykład

Macierz układów równań (70) i (72) ma następujący rozkład względem wartości osobliwych:

$$\begin{bmatrix} \frac{7}{6} & -\frac{1}{3} & -\frac{5}{6} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{5}{6} & -\frac{1}{3} & \frac{7}{6} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}. \quad (81)$$

Widzimy, że wektor $[1, 1, 1]^T$ rozpina jądro tej macierzy, natomiast kombinacje liniowe wektorów $[-1, 2, -1]^T$, $[-1, 0, 1]^T$ stanowią jej zasięg.

Ponieważ rozważana macierz jest symetryczna i rzeczywista (i nieujemnie określona), jej rozkład *SVD* jest zarazem jej rozkładem diagonalnym (w przypadku ogólnym taka zgodność nie zachodzi).

SVD i odwrotność macierzy

Niech $\mathbf{A} \in \mathbb{R}^{N \times N}$. Zauważmy, że $|\det \mathbf{A}| = \prod_{i=1}^N w_i$, a zatem $\det \mathbf{A} = 0$ wtedy i tylko wtedy, gdy co najmniej jeden $w_i = 0$. Niech $\det \mathbf{A} \neq 0$. Wówczas równanie $\mathbf{A}\mathbf{x} = \mathbf{b}$ ma rozwiązanie postaci

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} = \mathbf{V} [\text{diag}(w_i^{-1})] \mathbf{U}^T \mathbf{b}. \quad (82)$$

Niech teraz $\det \mathbf{A} = 0$. Równanie $\mathbf{A}\mathbf{x} = \mathbf{b}$ *także* ma rozwiązanie, o ile tylko $\mathbf{b} \in \text{Range } \mathbf{A}$. Rozwiązanie to ma postać $\mathbf{x} = \tilde{\mathbf{A}}^{-1}\mathbf{b}$, gdzie

$$\tilde{\mathbf{A}}^{-1} = \mathbf{V} [\text{diag}(\tilde{w}_i^{-1})] \mathbf{U}^T. \quad (83a)$$

gdzie

$$\tilde{w}_i^{-1} = \begin{cases} w_i^{-1} & \text{gdy } w_i \neq 0, \\ 0 & \text{gdy } w_i = 0. \end{cases} \quad (83b)$$

SVD i macierze osobliwe

Wróćmy jeszcze raz do problemu osobliwych (z zerowym wyznacznikiem głównym) układów równań, wspomnianego już na stronie 74. Jeżeli $\det \mathbf{A} = 0$, układ równań z całą pewnością nie ma *jednoznacznego* rozwiązania. Może jednak mieć rozwiązanie (a nawet nieskończenie wiele rozwiązań), jeżeli prawa strona *należy do zasięgu* \mathbf{A} . Jest to równoważne warunkowi, że wszystkie wyznaczniki poboczne we wzorach Cramera zerują się. Wówczas **rozwiązaniem** układu równań jest każdy wektor postaci

$$\mathbf{x} = \tilde{\mathbf{A}}^{-1} \mathbf{b} + \mathbf{x}_0, \quad (84)$$

gdzie $\tilde{\mathbf{A}}^{-1}$ jest pseudoodwrotnością daną przez (83), zaś $\mathbf{x}_0 \in \text{Ker} \mathbf{A}$ jest dowolnym wektorem należącym do jądra. Rozwiązanie z $\mathbf{x}_0 = \mathbf{0}$ ma spośród nich najmniejszą normę. Zauważmy, że na wektory należące do zasięgu, pseudoodwrotność działa jak zwykła odwrotność macierzy.

Jeżeli b *nie* należy do zasięgu, wyrażenie (84) z $x_0 = 0$ daje rozwiązanie przybliżone i najlepsze w sensie najmniejszych kwadratów, co niekiedy jest bardzo użyteczne.

Przykład

Zastosowanie pseudoodwrotności (83) do układu (72) daje

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}. \quad (85)$$

W tym wypadku pseudoodwrotność zachowuje się jak odwrotność (prawa strona równania (72) należy do zasięgu) i rozwiązanie (85) jest ścisłe.

Do rozwiązania (85) można dodać dowolny wektor postaci $[\alpha, \alpha, \alpha]$, a więc należący do jądra.

SVD i współczynnik uwarunkowania

Twierdzenie 2. Jeżeli macierz $\mathbf{A} \in \mathbb{R}^{N \times N}$ posiada rozkład (73) oraz $\det \mathbf{A} \neq 0$, jej współczynnik uwarunkowania spełnia

$$\kappa = \frac{\max_i |w_i|}{\min_i |w_i|}. \quad (86)$$

Jeśli macierz jest źle uwarunkowana, ale *formalnie* odwracalna, numeryczne rozwiązanie równania $\mathbf{A}\mathbf{x} = \mathbf{b}$ może być zdominowane przez wzmocniony błąd zaokrąglenia. Aby tego uniknąć, często zamiast (bezużytecznego!) rozwiązania dokładnego (82), używa się *przybliżonego* (i użytecznego!) rozwiązania w postaci (83) z następującą modyfikacją

$$\tilde{w}_i^{-1} = \begin{cases} w_i^{-1} & \text{gdy } |w_i| > \tau, \\ 0 & \text{gdy } |w_i| \leq \tau, \end{cases} \quad (87)$$

gdzie τ jest pewną zadaną tolerancją.

Przykład

Mamy rozwiązać następujące dwa układy równań:

$$\begin{bmatrix} 0.666666667 & -0.166666666 & -0.333333333 \\ -0.166666666 & 0.166666667 & -0.166666666 \\ -0.333333333 & -0.166666666 & 0.666666667 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1.284457048 \\ -0.577350273 \\ -0.129756514 \end{bmatrix} \quad (88a)$$

$$\begin{bmatrix} 0.666666667 & -0.166666666 & -0.333333333 \\ -0.166666666 & 0.166666667 & -0.166666666 \\ -0.333333333 & -0.166666666 & 0.666666667 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1.284457052 \\ -0.577350265 \\ -0.129756510 \end{bmatrix} \quad (88b)$$

Równania te różnią się tylko prawymi stronami, przy czym norma różnicy prawych stron jest rzędu 10^{-8} . Błąd takich rozmiarów łatwo może pojawić się w wyniku jakichś poprzednich obliczeń lub na skutek niepewności danych “zewnętrznych”, z którymi pracujemy.

Macierz w układach równań (88) jest symetryczna i dodatnio określona, a jej czynnik Cholesky’ego wynosi

$$\begin{bmatrix} 0.816496581 & & \\ -0.204124145 & 0.353553392 & \\ -0.408248290 & -0.707106777 & 0.000077460 \end{bmatrix} \quad (89)$$

Rozwiązania równań (88) za pomocą faktoryzacji Cholesky'ego mają postać

$$\mathbf{x}_a = \begin{bmatrix} -0.179434106 \\ -5.237183389 \\ -1.593647668 \end{bmatrix}, \quad \mathbf{x}_b = \begin{bmatrix} 3.903048668 \\ 2.927782158 \\ 2.488835105 \end{bmatrix}. \quad (90)$$

Różnica rozwiązań jest, co do normy, rzędu 10, czyli jest rzędu 10^9 razy większa, niż różnica prawych stron.

Faktoryzacja *SVD* macierzy z układów (88) pokazuje, że wartości szczególne tej macierzy są w przybliżeniu równe $1, \frac{1}{2}, 10^{-9}$. Jeśli do rozwiązania układów równań (88) zastosować pseudoodwrotność (87) ($\tau = 10^{-8}$), w obu wypadkach otrzymamy

$$\mathbf{x} = \begin{bmatrix} 1.861807320 \\ -1.154700538 \\ 0.447593757 \end{bmatrix}. \quad (91)$$

(91) jest jedynie *przybliżonym* rozwiązaniem równań (88). Jest ono jednak bardziej użyteczne, niż “ściśle” rozwiązania (90). Te dwa ostatnie najwyraźniej są zdominowane przez błąd, jaki wystąpił wzdłuż kierunku odpowiadającego najmniejszej wartości szczególnej macierzy. Nie wiemy — i nie mamy sposobu, aby to stwierdzić — które z dwu rozwiązań (90) jest “poprawne”. Przybliżone rozwiązanie (91) po prostu ignoruje wpływ tego kierunku, a więc i zaburzeń wzdłuż niego występujących.

Zauważmy, że jeśli podstawimy rozwiązanie (91) do lewych stron równań (88), w obu wypadkach otrzymamy

$$\begin{bmatrix} 1.2844571 \\ -0.5773503 \\ -0.12975654 \end{bmatrix} \quad (92)$$

Norma różnicy tego wektora i *obu* prawych stron równań (88) wynosi w przybliżeniu $6.664 \cdot 10^{-8}$.

Nadokreślone układy równań

Niech $\mathbf{A} \in \mathbb{R}^{M \times N}$, $M > N$, $\mathbf{b} \in \mathbb{R}^M$, $\mathbf{x} \in \mathbb{R}^N$. Wówczas układ równań

$$\mathbf{Ax} = \mathbf{b} \quad (93)$$

ma więcej równań, niż niewiadomych. Układ taki, zwany nadokreślonym, w ogólności nie ma rozwiązań. Za pomocą SVD można jednak znaleźć jego rozwiązanie przybliżone. Mianowicie

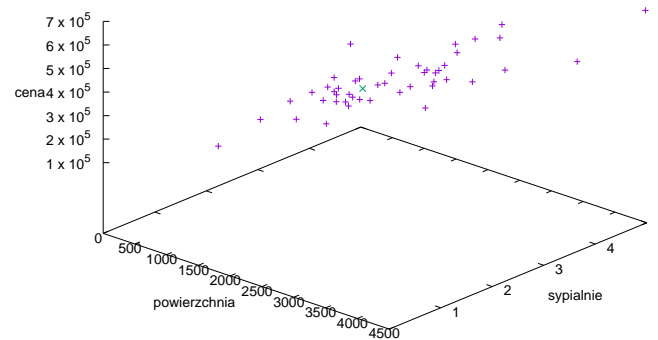
$$\|\mathbf{A} (\tilde{\mathbf{A}}^{-1} \mathbf{b}) - \mathbf{b}\|_2 = \text{minimum}, \quad (94)$$

gdzie $\tilde{\mathbf{A}}^{-1}$ jest pseudoodwrotnością (83). Widzimy, że $\tilde{\mathbf{A}}^{-1} \mathbf{b}$ jest przybliżonym, najlepszym w sensie najmniejszych kwadratów rozwiązaniem układu (93). Metoda ta jest *powszechnie* używana w liniowym zagadnieniu najmniejszych kwadratów.

Przykład

Ceny domów w Portland*.

W pliku znajdują się dane dotyczące cen domów. Cena zależy od powierzchni użytkowej domu i od liczby sypialni. Mamy do dyspozycji 47 danych. Chcemy się dowiedzieć, ile *powinna* wynosić cena domu o powierzchni 1650 sq. ft, z trzema sypialniami. Rysunek przedstawia dostępne dane:



*Przykład pochodzi z kursu Machine Learning na Uniwersytecie Stanforda. Dane są autentyczne.

Widzimy, że dane *w przybliżeniu* układają się na pewnej płaszczyźnie. Proponujemy wobec tego model:

$$\text{cena} = \theta_0 + \text{powierzchnia} \cdot \theta_1 + \text{sypialnie} \cdot \theta_2 \quad (95)$$

Równania tego typu powinny być spełnione dla *wszystkich* dostępnych danych:

$$\theta_0 + p_1\theta_1 + s_1\theta_2 = c_1 \quad (96a)$$

$$\theta_0 + p_2\theta_1 + s_2\theta_2 = c_2 \quad (96b)$$

⋮

$$\theta_0 + p_{47}\theta_1 + s_{47}\theta_2 = c_{47} \quad (96c)$$

Jest to układ 47 równań na trzy niewiadome $(\theta_0, \theta_1, \theta_2)$ i jest niezwykle mało prawdopodobne, że uda się go rozwiązać ściśle, to znaczy tak, aby

każde z równań (96a) było ściśle spełnione. Musimy zadowolić się rozwiązaniem przybliżonym.

W postaci stabelaryzowanej dane mają postać

Nr	powierznia	sypialnie	cena
1	2100	3	399900
2	1600	3	329900
3	2400	3	369000
.....			
45	852	2	179900
46	1852	4	299900
47	1203	3	239500

Uwaga: Wśród danych wejściowych, jedna (powierzchnia) przybiera wartości rzędu tysięcy, druga (sypialnie) rzędu jedności. Może to, potencjalnie, być źródłem błędów numerycznych. Dlatego dane warto *znormalizować*, to znaczy odjąć od nich wartości średnie poszczególnych wielkości i podzielić przez odchylenia standardowe.

Równania (96a) zapisujemy w postaci macierzowej:

$$\mathbf{X} \cdot \boldsymbol{\theta} = \mathbf{c}, \quad (97)$$

gdzie

$$\mathbf{X} = \begin{bmatrix} 1 & p_1 & s_1 \\ 1 & p_2 & s_2 \\ 1 & p_3 & s_3 \\ \dots & \dots & \dots \\ 1 & p_{45} & s_{45} \\ 1 & p_{46} & s_{46} \\ 1 & p_{47} & s_{47} \end{bmatrix}$$

$\boldsymbol{\theta} = [\theta_0, \theta_1, \theta_2]^T$, $\mathbf{c} = [c_1, c_2, c_3, \dots, c_{45}, c_{46}, c_{47}]^T$, a p_i, s_i, c_i są (znormalizowaną) powierzchnią, liczbą sypialni i ceną i -tego domu. \mathbf{X} jest macierzą o 47 wierszach i 3 kolumnach; pierwsza kolumna macierzy \mathbf{X} zawiera same jedynki.

Korzystając z naszego ulubionego pakietu obliczającego faktoryzację *SVD* macierzy \mathbf{X} , do równania (97) stosujemy pseudoodwrotność (83) i znajdujemy *optymalne* rozwiązanie przybliżone

$$\boldsymbol{\theta}_{\text{opt}} = \tilde{\mathbf{X}}^{-1} \mathbf{c} \quad (98)$$

Następnie podstawiamy (znormalizowane) dane wejściowe ($p = 1650, s = 3$) i (po denormalizacji) uzyskujemy odpowiedź, iż taki dom powinien kosztować $\boldsymbol{\theta}_{\text{opt}} \cdot [1, p, s]^T = \$ 293081.46$.