Metody Numeryczne

Praca domowa 1.

Hubert Michalski hm438596

współpraca: Przemysław Fuchs

6 lutego 2024

Zadanie 1.2

Niech $A \in R^{N \times N}$ będzie nieosobliwą macierzą trójdiagonalną. Podaj algorytm, który przy użyciu przekształceń Householdera wyznaczy jej rozkład QR możliwie niskim (jakim?) kosztem. Wskazówka: Być może macierz Q warto wyznaczyć w postaci iloczynu pewnych przekształceń.

Rozwiązanie

Zapiszmy najpierw standardowy algorytm rozkładu QR za pomocą przekształceń Householdera:

Algorithm 1 $A = QR$ Householder method (iterative version)

1: **for** k $= 1 : N$ **do**

2:     Podziel $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & A_{22} \end{bmatrix}$, $a_{11} \in \mathbb{R}$

3:     $Q_k =$ m. Householdera t.że $H \cdot \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} = \begin{bmatrix} r_{kk} \\ 0 \end{bmatrix}$

4:     $\begin{bmatrix} r_{12} \\ B_{22} \end{bmatrix} = Q_k^T \cdot \begin{bmatrix} a_{12} \\ A_{22} \end{bmatrix}$

5:     $A = B_{22}$

6: **return** $Q = Q_1 \begin{bmatrix} 1 & \\ & Q_2 \end{bmatrix} \cdots \begin{bmatrix} I_{k-1} & \\ & Q_k \end{bmatrix} \cdots, R = \begin{bmatrix} r_{11} & r_{12} \\ 0 & \ddots \end{bmatrix}$

Jeśli pokażemy, że kroki 3. i 4. da się wykonać w czasie stałym, to cały algorytm rozkładu $A = QR$ będzie miał złożoność liniową. Jako że na wejściu otrzymujemy liniowo wiele zmiennych, gdzie w oczywisty sposób każda wpływa na wynik, to jest to także ograniczenie dolne i nie da się przedstawić szybszego algorytmu.

Udowodnimy najpierw, że macierz Householdera taką, jak w kroku 3. da się wyznaczyć w $\mathcal{O}(1)$. Pamiętamy, że m. Householdera można reprezentować za pomocą wektora v takiego, że $H = I - \gamma v v^T$ gdzie $\gamma = 2/||v||_2^2$. Zauważmy, że dany wektor:

$$\vec{a} = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}$$

ma jedynie dwie niezerowe wartości, ponieważ macierz A jest trójdiagonalna. Potem dokładnie pokażemy, że ta własność macierzy jest zachowana w każdej iteracji. Żeby wyznaczyć wektor v możemy skorzystać ze wzoru $v = \vec{a} + sgn(a_{11})||\vec{a}||_2 \vec{e_1}$, gdzie policzenie normy to koszt $\mathcal{O}(2)$ ponieważ tylko pierwsze dwa elementy są niezerowe, co ostatecznie daje nam koszt stały dla tego kroku algorytmu. Dodatkowo zwróćmy uwagę na to, że nie przedstawiamy macierzy Q_k dokładnie (tzn. jako faktycznie macierzy) a jedynie jako wektor który ma dwa elementy lub pamiętamy go jako dwie stałe.

Aby udowodnić, że 4. krok algorytmu da się wykonać w czasie stałym, przyjrzyjmy się najpierw postaci macierzy Q_k:

$$Q_k = I - \gamma v v^T = \begin{bmatrix} 1 & & \dots & 0 \\ & 1 & & \vdots \\ \vdots & & \ddots & \\ 0 & \dots & & 1 \end{bmatrix} - \gamma \begin{bmatrix} v_1 \\ v_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{bmatrix} v_1 & v_2 & 0 & \dots & 0 \end{bmatrix} = \begin{bmatrix} 1-\gamma v_1^2 & -\gamma v_1 v_2 & \dots & 0 \\ -\gamma v_1 v_2 & 1-\gamma v_2^2 & & \vdots \\ \vdots & & \ddots & \\ 0 & \dots & \dots & 1 \end{bmatrix}$$

Widzimy, że po przemnożeniu z lewej przez Q_k jedynie dwa pierwsze wiersze macierzy zostaną zmienione, co już potencjalnie zmniejsza ilość obliczeń potrzebnych do uzyskania wyniku. Możemy jednak posunąć się jeszcze o krok dalej i przeanalizować jakiej postaci jest cały iloczyn:

$$\begin{bmatrix} r_{12} \\ B_{22} \end{bmatrix} = Q_k^T \cdot \begin{bmatrix} a_{12} \\ \hdashline A_{22} \end{bmatrix} = \begin{bmatrix} 1-\gamma v_1^2 & -\gamma v_1 v_2 & \dots & 0 \\ -\gamma v_1 v_2 & 1-\gamma v_2^2 & & \vdots \\ \vdots & & \ddots & \\ 0 & \dots & \dots & 1 \end{bmatrix} \begin{bmatrix} \tau & 0 & \dots & 0 \\ \hdashline a'_{11} & a'_{12} & & \\ a'_{21} & a'_{22} & a'_{23} & \\ & a'_{32} & a'_{33} & \ddots \\ & & \ddots & \ddots \end{bmatrix}$$

gdzie a'_{ij} to elementy macierzy A_{22}, a τ to jedyny element wektora a_{12} z przedstawionego uprzednio pseudokodu. Łatwo teraz zauważyć, że obliczenia jakie trzeba wykonać to:

$$r'_1 = (1 - \gamma v_1^2) \cdot \tau - \gamma v_1 v_2 \cdot a'_{11}$$

$$r'_2 = -\gamma v_1 v_2 \cdot a'_{12}$$

więc wynikowy wektor to

$$r_{12} = \begin{bmatrix} r'_1 & r'_2 & 0 & \dots & 0 \end{bmatrix}$$

następnie liczymy pierwszy wiersz macierzy B_{22}:

$$b_{11} = (-\gamma v_1 v_2) \cdot \tau + (1 - \gamma v_2^2) \cdot a'_{11}$$

$$b_{12} = (1 - \gamma v_2^2) \cdot a'_{12}$$

Zatem jesteśmy w stanie wykonać iloczyn zadany w kroku 4. w czasie stałym, ponieważ nie przepisujemy całej macierzy do B_{22} a jedynie nadpisujemy pierwszy wiersz macierzy A_{22}. Warto także zauważyć, ze macierz B_{22} także będzie diagonalna zatem kolejne iteracje będą przebiegały analogicznie. Dodatkowo wynikowy wektor r_{12} posiada jedynie dwie stałe, jest to istotne ponieważ w każdej iteracji do macierzy wynikowej R dokładamy jedynie stałą ilość zmiennych więc będzie ich liniowo wiele względem wejścia.

Podsumowując, jesteśmy w stanie poznać rozkład $A = QR$ dla macierzy trójdiagonalnej nieosobliwej w czasie $\mathcal{O}(N)$ jeśli będziemy reprezentować macierze Householdera Q_k za pomocą ich wektorów v_k, gdzie będziemy pamiętać jedynie niezerowe wartości. Analogicznie dla macierzy R - pamiętamy jedynie niezerowe wartości w wierszach a nie całe wektory.