Householder-transzformáció

2016. október 23.

Tekintsünk egy v egység hosszúságú n dimenziós vektort, azaz $\|v\|_2 = 1$. A vektorhoz tartozó Householder-transzformációt a következőképpen definiáljuk.

$$H(v) := I - 2vv^T$$

ahol I az n dimenziós egységmátrix. Mivel v egy oszlopvektor, v^T egy sorvektor, így vv^T egy $n \times n$-es mátrix. A Householder-mátrixszal dolgozni egyszerű, ugyanis nem kell előállítani magát a mátrixot ahhoz, hogy számolni tudjunk vele. Legyen a egy tetszőleges n-dimenziós vektor, és vizsgáljuk a $H(v)a$ mátrix-vektor szorzást

$$H(v)a = I - 2(vv^T)a = a - 2v(v^T a) = a - 2v\langle v, a\rangle$$

Felhasználva tehát az asszociativitást a mátrix-vektor szorzás kiváltható egy skalárszorzat kiszámításával, így tehát csak vektorműveleteket kell végeznünk, hogy meghatározzuk a transzformáció hatását egy tetszőleges a vektorra.
A Householder-transzformáció a v vektor ortogonális kiegészítő alterére való tükrözés. Ez két dimenzióban a v-re merőleges egyenesre tükrözés, három dimenzióban a v normálisú síkra vett tükrözés. Ebből következik, hogy ha v-t az a-ból b-be mutató egység hosszúságú vektornak választjuk, akkor az ehhez tartozó Householder-transzformáció az a vektort a b-re a b-t pedig az a-ra képezi, azaz

$$v = \frac{a-b}{\|a-b\|} \quad \Rightarrow \quad H(v)a = b, \quad H(v)b = a$$

Ez pedig azt jelenti, hogy a Householder-mátrixszal egy tetszőleges a vektort áttranszformálhatunk egy általunk választott alakú b vektorra. Legyen most $b = \sigma e_1$, ahol e_1 az első kanonikus egységvektor, és $\sigma = -sign(a_1)\|a\|_2$. Ekkor

$$v = \frac{a - \sigma e_1}{\|a - \sigma e_1\|_2} \quad \Rightarrow \quad H(v)a = \sigma e_1$$

A Householder-transzformációval tehát ki tudjuk nullázni egy $n \times n$ mátrix első oszlopának $n-1$ elemét. Az első transzformáció után kapott mátrix $n-1 \times n-1$ méretű részmátrixának hasonló transzformációjával kinullázhatjuk a második oszlop $n-2$ elemét. Folytatva ezt az eljárást, végül felső háromszög alakra hozhatjuk a

mátrixot. A Householder transzformációkhoz tartozó mátrixok ortogonálisak, ezért ezek szorzata és inverze is ortogonális, végső soron a transzformációkhoz tartozó mátrixok szorzata adja az $A = QR$ felbontásban szereplő Q ortogonális mátrixot.
3×3 méretű mátrixok esetén a következőképpen járhatunk el. Legyen

$$A = (a_1 \ a_2 \ a_3) = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Az első Householder-mátrixot az előzőek szerint így kapjuk:

$$\sigma = -sign(a_{11})\sqrt{a_{11}^2 + a_{21}^2 + a_{31}^2}$$

$$v_1 = \frac{a_1 - \sigma e_1}{\|a_1 - \sigma e_1\|}$$

$$H_1 := H(v_1) = I - 2v_1 v_1^T$$

A transzformációt alkalmazva az A mátrix egyes oszlopaira a következőt kapjuk:

$$H_1 A = (H(v_1)a_1 \ \ H(v_1)a_2 \ \ H(v_1)a_3) = (\sigma_1 e_1 \ \ \tilde{a}_2 \ \ \tilde{a}_3) = \begin{pmatrix} \sigma_1 & \tilde{a}_{12} & \tilde{a}_{13} \\ 0 & \tilde{a}_{22} & \tilde{a}_{23} \\ 0 & \tilde{a}_{32} & \tilde{a}_{33} \end{pmatrix}$$

A következő transzformációt úgy kapjuk, hogy a jobb alsó 2×2-es részmátrix Householder-transzformációját írjuk fel. Legyen

$$\sigma_2 = -sign(\tilde{a}_{22})\sqrt{\tilde{a}_{22}^2 + \tilde{a}_{32}^2}$$

Továbbá

$$v_2 = \frac{\begin{pmatrix} \tilde{a}_{22} \\ \tilde{a}_{32} \end{pmatrix} - \begin{pmatrix} \sigma_2 \\ 0 \end{pmatrix}}{\left\| \begin{pmatrix} \tilde{a}_{22} \\ \tilde{a}_{32} \end{pmatrix} - \begin{pmatrix} \sigma_2 \\ 0 \end{pmatrix} \right\|_2}$$

A $H(v_2) = I - 2v_2 v_2^T$ mátrix most 2×2-es, ezért a jobb alsó sarokban lévő ugyanekkora részmátrix transzformációját írja le, ezért konstruálunk egy olyan 3×3-as mátrixot, amely helybenhagyja az első sor és első oszlop elemeit, a többi elemet pedig $H(v_2)$ szerint transzformálja. Így ezzel a mátrixszal már közvetlenül megszorozhatjuk az A mátrixot:

$$H_2 := \begin{pmatrix} 1 & 0 & 0 \\ 0 & H(v_2)_{11} & H(v_2)_{12} \\ 0 & H(v_2)_{21} & H(v_2)_{22} \end{pmatrix}$$

Végül a transzformáció hatása a H_1A mátrixra:

$$H_2H_1A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & H(v_2)_{11} & H(v_2)_{12} \\ 0 & H(v_2)_{21} & H(v_2)_{22} \end{pmatrix} \begin{pmatrix} \sigma_1 & \tilde{a}_{12} & \tilde{a}_{13} \\ 0 & \tilde{a}_{22} & \tilde{a}_{23} \\ 0 & \tilde{a}_{32} & \tilde{a}_{33} \end{pmatrix} = \begin{pmatrix} \sigma_1 & \tilde{a}_{12} & \tilde{a}_{13} \\ 0 & \sigma_2 & \tilde{\tilde{a}}_{23} \\ 0 & 0 & \tilde{\tilde{a}}_{33} \end{pmatrix} =: R$$

Az így kapott R mátrix már felső háromszög alakú, továbbá mivel H_2 és H_1 ortogonális mátrixok, megkaptuk az $A = QR$ felbontást is.

$$A = (H_2H_1)^{-1}R = H_1^{-1}H_2^{-1}R = H_1^TH_2^TR = QR$$

Példaképpen hozzuk felső háromszög alakra a következő mátrixot:

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

Az előző jelölésekkel:

$$\sigma_1 = -sign(1)\sqrt{1^2 + 2^2 + 2^2} = -3$$

$$a_1 - \sigma_1 e_1 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} - \begin{pmatrix} -3 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 2 \end{pmatrix}$$

$$v_1 = \frac{a_1 - \sigma_1 e_1}{\|a_1 - \sigma_1 e_1\|_2} = \frac{1}{\sqrt{4^2 + 2^2 + 2^2}} \begin{pmatrix} 4 \\ 2 \\ 2 \end{pmatrix} = \frac{1}{\sqrt{24}} \begin{pmatrix} 4 \\ 2 \\ 2 \end{pmatrix}$$

A transzformációt oszloponként végezzük el. Az első oszlopot csak ellenőrzésképpen számoljuk ki, ezt ugye tudjuk, hogy milyen alakú

$$H(v_1)a_1 = a_1 - 2v_1(v_1^Ta_1) = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} - 2 \cdot \frac{1}{\sqrt{24}} \begin{pmatrix} 4 \\ 2 \\ 2 \end{pmatrix} \frac{1}{\sqrt{24}} \begin{pmatrix} 4 & 2 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} =$$

$$\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} - \frac{2}{24} \begin{pmatrix} 4 \\ 2 \\ 2 \end{pmatrix} (4 \cdot 1 + 2 \cdot 2 + 2 \cdot 2) = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} - \begin{pmatrix} 4 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} -3 \\ 0 \\ 0 \end{pmatrix}$$

$$H(v_1)a_2 = a_2 - 2v_1(v_1^Ta_2) = \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} - 2 \cdot \frac{1}{\sqrt{24}} \begin{pmatrix} 4 \\ 2 \\ 2 \end{pmatrix} \frac{1}{\sqrt{24}} \begin{pmatrix} 4 & 2 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} =$$

$$\begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} - \frac{2}{24} \begin{pmatrix} 4 \\ 2 \\ 2 \end{pmatrix} (4 \cdot 0 + 2 \cdot 0 + 2 \cdot 3) = \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} - \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix}$$

$$H(v_1)a_3 = a_3 - 2v_1(v_1^T a_3) = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} - 2 \cdot \frac{1}{\sqrt{24}} \begin{pmatrix} 4 \\ 2 \\ 2 \end{pmatrix} \frac{1}{\sqrt{24}} \begin{pmatrix} 4 & 2 & 2 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} =$$

$$\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} - \frac{2}{24} \begin{pmatrix} 4 \\ 2 \\ 2 \end{pmatrix} (4 \cdot (-1) + 2 \cdot 1 + 2 \cdot 1) = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

Így a transzformált mátrix a következő:

$$H_1 A = \begin{pmatrix} -3 & -2 & -1 \\ 0 & -1 & 1 \\ 0 & 2 & 1 \end{pmatrix}$$

Tekintsük most a hátralévő 2×2-es részmátrixot.

$$\sigma_2 = -sign(-1)\sqrt{(-1)^2 + 2^2} = \sqrt{5}$$

$$v_2 = \frac{\begin{pmatrix} -1 \\ 2 \end{pmatrix} - \begin{pmatrix} \sqrt{5} \\ 0 \end{pmatrix}}{\left\| \begin{pmatrix} -1 \\ 2 \end{pmatrix} - \begin{pmatrix} \sqrt{5} \\ 0 \end{pmatrix} \right\|_2} = \frac{\begin{pmatrix} -1 - \sqrt{5} \\ 2 \end{pmatrix}}{\sqrt{(-1 - \sqrt{5})^2 + 2^2}} = \frac{1}{\sqrt{1 + 2\sqrt{5} + 5 + 4}} \begin{pmatrix} -1 - \sqrt{5} \\ 1 \end{pmatrix} =$$

$$\frac{1}{\sqrt{10 + 2\sqrt{5}}} \begin{pmatrix} -1 - \sqrt{5} \\ 2 \end{pmatrix}$$

A rész oszlopokon végezzük el a transzformációt

$$H(v_2)\tilde{a}_2 = \begin{pmatrix} -1 \\ 2 \end{pmatrix} - 2 \cdot \frac{1}{\sqrt{10 + 2\sqrt{5}}} \begin{pmatrix} -1 - \sqrt{5} \\ 2 \end{pmatrix} \frac{1}{\sqrt{10 + 2\sqrt{5}}} \begin{pmatrix} -1 - \sqrt{5} & 2 \end{pmatrix} \begin{pmatrix} -1 \\ 2 \end{pmatrix} =$$

$$\begin{pmatrix} -1 \\ 2 \end{pmatrix} - \frac{2}{10 + 2\sqrt{5}} \begin{pmatrix} -1 - \sqrt{5} \\ 2 \end{pmatrix} ((-1 - \sqrt{5})(-1) + 2 \cdot 2) =$$

$$\begin{pmatrix} -1 \\ 2 \end{pmatrix} - \frac{2(5 + \sqrt{5})}{10 + 2\sqrt{5}} \begin{pmatrix} -1 - \sqrt{5} \\ 2 \end{pmatrix} =$$

$$\begin{pmatrix} -1 \\ 2 \end{pmatrix} - \frac{10 + 2\sqrt{5}}{10 + 2\sqrt{5}} \begin{pmatrix} -1 - \sqrt{5} \\ 2 \end{pmatrix} = \begin{pmatrix} \sqrt{5} \\ 0 \end{pmatrix}$$

$$H(v_2)\tilde{a}_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - 2 \cdot \frac{1}{\sqrt{10+2\sqrt{5}}} \begin{pmatrix} -1-\sqrt{5} \\ 2 \end{pmatrix} \frac{1}{\sqrt{10+2\sqrt{5}}} \begin{pmatrix} -1-\sqrt{5} & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} =$$

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \frac{2}{10+2\sqrt{5}} \begin{pmatrix} -1-\sqrt{5} \\ 2 \end{pmatrix} ((-1-\sqrt{5}) \cdot 1 + 2 \cdot 1) =$$

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \frac{2(1-\sqrt{5})}{10+2\sqrt{5}} \begin{pmatrix} -1-\sqrt{5} \\ 2 \end{pmatrix} =$$

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{1-\sqrt{5}}{5+\sqrt{5}} \begin{pmatrix} 1+\sqrt{5} \\ -2 \end{pmatrix} = \frac{1}{5+\sqrt{5}} \begin{pmatrix} 5+\sqrt{5}+1-5 \\ 5+\sqrt{5}-2+2\sqrt{5} \end{pmatrix} = \frac{1}{5+\sqrt{5}} \begin{pmatrix} 1+\sqrt{5} \\ 3+3\sqrt{5} \end{pmatrix}$$

A transzformált mátrix alakja pedig a következő:

$$H_2 H_1 A = R = \begin{pmatrix} -3 & -2 & -1 \\ 0 & \sqrt{5} & \frac{1+\sqrt{5}}{5+\sqrt{5}} \\ 0 & 0 & \frac{3+3\sqrt{5}}{5+\sqrt{5}} \end{pmatrix}$$