Numerikus módszerek 1.

5. előadás: QR-felbontás: Gram–Schmidt ortogonalizáció, Householder-transzformációk és alkalmazásaik

Krebsz Anna

ELTE IK

Tartalomjegyzék

Tartalomjegyzék

Definíció: ortogonális mátrix

Egy $Q \in \mathbb{R}^{n\times n}$ mátrix *ortogonális*, ha az inverze a transzponáltja, azaz

$$Q^\top Q = I.$$

Megj.: Ekkor $QQ^\top = I$ is teljesül. $(Q^{-1} = Q^\top)$

Definíció: skaláris szorzat

Az $x, y \in \mathbb{R}^n$ vektorok *skaláris szorzata*

$$\langle x, y\rangle := y^\top x = \sum_{k=1}^{n} x_k \cdot y_k.$$

Definíció: ortonormált rendszer

A $q_1, \ldots, q_n \in \mathbb{R}^n$ vektorok *ortonormált rendszert* alkotnak, ha

$$\langle q_i, q_j \rangle = \begin{cases} 0 & \text{ha } i \neq j, \\ 1 & \text{ha } i = j. \end{cases}$$

Állítás: ortogonális mátrixok oszlopvektorairól

A $Q \in \mathbb{R}^{n \times n}$ ortogonális mátrix oszlopai, mint vektorok ortonormált rendszert alkotnak.

Biz.: Gondoljunk bele: $Q^\top Q = I$. $\square$

Definíció: ortogonális rendszer

A $q_1, \ldots, q_n \in \mathbb{R}^n$ vektorok *ortogonális rendszert* alkotnak, ha

$$\langle q_i, q_j \rangle = 0 \qquad (i \neq j).$$

Állítás: ortogonális rendszerekből álló mátrixokról

Ha a $q_1, \ldots, q_n \in \mathbb{R}^n$ vektorok ortogonális rendszert alkotnak, akkor a $Q := (q_1, \ldots, q_n) \in \mathbb{R}^{n \times n}$ mátrix esetén a $Q^\top Q$ szorzatmátrix diagonális. ($QQ^\top$ általában nem.)

Biz.: Gondoljunk bele: $Q^\top Q = D$ diagonális mátrix. $\square$

Elnevezések:

- $\langle q_i, q_j \rangle = \delta_{ij}$ (Kronecker-féle delta).
- $q_i \perp q_j \Leftrightarrow \langle q_i, q_j \rangle = 0 \ \ (i \neq j)$: az oszlopok merőlegesek, avagy *ortogonálisak* egymásra
- $\langle q_i, q_i \rangle = 1$: minden oszlopvektor hossza 1, avagy *normált* $\|q_i\|_2 := \sqrt{\langle q_i, q_i \rangle}$: „hossz”, avagy „kettes norma”

Példa: ortogonális mátrixok

Az alábbi mátrixok ortogonálisak:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \qquad \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

Állítás: ortogonális mátrixok szorzata

Ha $Q_1, Q_2 \in \mathbb{R}^{n\times n}$ ortogonális mátrixok, akkor a szorzatuk, $Q_1 Q_2$ is ortogonális.

Biz.: Tudjuk, hogy $Q_1^\top Q_1 = I$ és $Q_2^\top Q_2 = I$.

Kell, hogy $Q_1 Q_2$ is ortogonális.

Vizsgáljuk:

$$(Q_1 Q_2)^\top (Q_1 Q_2) = Q_2^\top \underbrace{Q_1^\top Q_1}_{I} Q_2 = Q_2^\top Q_2 = I.$$

$\square$

Definíció: QR-felbontás

Az $A \in \mathbb{R}^{n \times n}$ mátrix QR-felbontásának nevezzük a $Q \cdot R$ szorzatot, ha $A = QR$, ahol $Q \in \mathbb{R}^{n \times n}$ ortogonális mátrix, $R \in \mathcal{U}$ pedig felső háromszögmátrix.

Tétel: QR-felbontás létezése és egyértelműsége

Ha $\det A \neq 0$, (vagyis az A oszlopvektorai lineárisan függetlenek), akkor A-nak létezik QR-felbontása.
Ha még feltesszük, hogy $r_{ii} > 0 \ \forall\, i$-re, akkor egyértelmű is.

Biz.: Létezés: A bizonyítást a Gram–Schmidt-féle ortogonalizációs eljárás adja: az A mátrix oszlopaiból – amelyek a feltétel értelmében lineárisan függetlenek – előállítjuk a Q oszlopait és R ismeretlen elemeit.

Tekintsük a $Q \cdot R = A$ mátrixszorzást, ahol A-t és Q-t az oszlopaival adtuk meg:

$$\begin{bmatrix} q_1 & q_2 & \dots & q_n \end{bmatrix} \cdot \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & \dots & r_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & r_{nn} \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix}.$$

Tekintsük először A első oszlopát, a_1-et. A mátrixszorzásból

$$r_{11} \cdot q_1 = a_1, \;\Rightarrow\; q_1 = \frac{1}{r_{11}} \cdot a_1.$$

Mivel q_1-től azt várjuk el, hogy normált legyen, ezért $r_{11} := \|a_1\|_2$.

QR-felbontás létezés bizonyítás

Tegyük fel, hogy A első $k-1$ oszlopát már felhasználtuk, és így előállítottuk Q első $k-1$ oszlopát, melyek normáltak és egymásra ortogonálisak, valamint R első $k-1$ oszlopának elemeit is ismerjük.

Tekintsük most a_k-t. A mátrixszorzásból felírhatjuk a_k-t, majd kifejezhetjük q_k-t:

$$a_k = \sum_{j=1}^{k} r_{jk} \cdot q_j \quad \Longrightarrow \quad q_k = \frac{1}{r_{kk}} \left(a_k - \sum_{j=1}^{k-1} r_{jk} \cdot q_j \right)$$

Az r_{jk} értékek meghatározásához szorozzuk be skalárisan mindkét oldalt q_i-vel rögzített i értékre $(i = 1, 2, \ldots, k-1)$ és használjuk ki, hogy $\langle q_i, q_j \rangle = \delta_{ij}$, valamint q_k-tól is azt várjuk, hogy merőleges legyen az összes eddigi q_i vektorra:

$$q_k = \frac{1}{r_{kk}} \left(a_k - \sum_{j=1}^{k-1} r_{jk} \cdot q_j \right) \quad \left| \cdot q_i \right\rangle \quad (i = 1, \ldots, k-1)$$

$$0 = \langle q_k, q_i \rangle = \frac{1}{r_{kk}} \left(\langle a_k, q_i \rangle - \sum_{j=1}^{k-1} r_{jk} \underbrace{\langle q_j, q_i \rangle}_{\delta_{ij}} \right) =$$

$$= \frac{1}{r_{kk}} \left(\langle a_k, q_i \rangle - r_{ik} \right) \quad \Rightarrow \quad r_{ik} = \langle a_k, q_i \rangle \,.$$

Továbbá q_k-tól még azt várjuk el, hogy normált legyen, ezért

$$r_{kk} = \left\| a_k - \sum_{j=1}^{k-1} r_{jk} \cdot q_j \right\|_2 \,.$$

Így megkaptuk az R mátrix k-adik oszlopának ismeretlen értékeit, az előállított q_k ortogonális az eddigi q_i-kre, valamint normált. $\square$

Biz.: Egyértelműség: Tegyük fel indirekt, hogy legalább két különböző QR-felbontásunk van

$$A = Q_1 R_1 = Q_2 R_2,$$

melyekre a R_1 és R_2 diagonális elemi pozitívak.

A-t szorozzuk balról $Q_2^{-1} = Q_2^\top$-tal és jobbról R_1^{-1}-zel

$$\underbrace{(Q_2^\top Q_1)}_{\text{ortogonális}} = \underbrace{(R_2 R_1^{-1})}_{\in\mathcal{U}}.$$

Legyen $R := R_2 R_1^{-1}$, mivel $Q := Q_2^\top Q_1$ ortogonális mátrix ($R = Q$),

$$Q^\top Q = I = R^\top R.$$

QR-felbontás egyértelműség bizonyítás

Az $R^\top R = I$ szorzatot felírva:

$$
\begin{array}{cc}
 & \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & 0 & \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & r_{nn} \end{bmatrix} \\
\begin{bmatrix} r_{11} & 0 & \dots & \\ r_{12} & r_{22} & 0 & \\ \vdots & & \ddots & \vdots \\ r_{1n} & 0 & \dots & r_{nn} \end{bmatrix} & \begin{bmatrix} 1 & 0 & \dots & \\ 0 & 1 & 0 & \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}
\end{array}
$$

$r_{11} \cdot r_{11} = 1$, amiből $r_{11} > 0$ miatt $r_{11} = 1$.
$j \neq 1$-re

$$r_{11} \cdot r_{1j} = 0 \quad \Rightarrow \quad r_{1j} = 0.$$

R második sorára: $r_{22} \cdot r_{22} = 1$, amiből $r_{22} > 0$ miatt $r_{22} = 1$.

A szorzat mátrix $(2, j)$-edik elemére $j \neq 2$-re

$$r_{22} \cdot r_{2j} = 0 \quad \Rightarrow \quad r_{2j} = 0.$$

A többi sorra ehhez hasonlóan ellenőrizhetjük, hogy

$$R = I \;\Leftrightarrow\; R_1 = R_2,\;\; Q_1 = Q_2.$$

Ezzel ellentmondásra jutottunk. $\square$

Megj.: Két különböző QR-felbontás esetén létezik olyan $D := \operatorname{diag}(\pm 1, \ldots, \pm 1)$ mátrix, melyre $A = \overbrace{Q \cdot D} \cdot \overbrace{D \cdot R} = \widetilde{Q} \cdot \widetilde{R}$.

Miért jó a QR-felbontás?

Tegyük fel, hogy

- az $Ax = b$ LER megoldható, és
- rendelkezésünkre áll az $A = QR$ felbontás.

Ekkor $Ax = Q \cdot \underbrace{R \cdot x}_{y} = b$ helyett $\quad (\frac{2}{3}n^3 + \mathcal{O}(n^2))$

1. a $Qy = b$ LER megoldása: $y = Q^\top b$, $\quad (2n^2 + \mathcal{O}(n))$
2. az $Rx = y$ LER-t oldjuk meg. $\quad (n^2 + \mathcal{O}(n))$

Együtt is írható: oldjuk meg az $Rx = Q^\top b$ LER-t.

Persze valamikor elő kell állítani a QR-felbontást. $\quad (2n^3 + \mathcal{O}(n^2))$
Előnyös, ha sokszor ugyanaz A, lásd QR-algoritmus (Num. mód. 2A). Így numerikusan stabilabb a LER megoldása.

Feladat: adott $a_1, \ldots, a_n \in \mathbb{R}^n$ lineárisan független vektorrendszer, készítsünk belőlük egy $q_1, \ldots, q_n \in \mathbb{R}^n$ ortonormált vektorrendszert úgy, hogy q_k csak $a_1, \ldots, a_k$-tól függ ($k = 1, 2, \ldots, n$).

Másképp, mátrixszorzás alakban: $QR = A$, avagy

$$\begin{bmatrix} q_1 & q_2 & \ldots & q_n \end{bmatrix} \cdot \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & \cdots & r_{2n} \\ & & \ddots & \vdots \\ & & & r_{nn} \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & \ldots & a_n \end{bmatrix}$$

Adott: A, keressük: Q, R.

Levezetés: lásd a QR-felbontás létezés bizonyítását (illetve Linalg).

Definíció: Gram–Schmidt-féle ortogonalizáció

Adott az $a_1, \ldots, a_n \in \mathbb{R}^n$ lineárisan független vektorrendszer.

1. $r_{11} := \|a_1\|_2$,
2. $q_1 := \dfrac{1}{r_{11}} a_1$ („lenormáljuk").

A k-adik lépésben $(k = 2, \ldots, n)$:

3. $r_{jk} := \langle a_k, q_j \rangle \quad (j = 1, \ldots, k-1)$,
4. $s_k := a_k - \sum_{j=1}^{k-1} r_{jk} \cdot q_j$,
5. $r_{kk} := \|s_k\|_2$ (s_k segédvektor hossza),
6. $q_k := \dfrac{1}{r_{kk}} s_k$ („lenormáljuk").

Az így nyert $q_1, \ldots, q_n \in \mathbb{R}^n$ vektorrendszer ortonormált.

Definíció: Gram–Schmidt-ortogonalizáció (normálás nélkül)

Adott az $a_1, \ldots, a_n \in \mathbb{R}^n$ lineárisan független vektorrendszer.

1. $\widetilde{q_1} := a_1$,
2. $\widetilde{r_{11}} := 1$

   A k-adik lépésben ($k = 2, \ldots, n$):
3. $\widetilde{r_{jk}} := \dfrac{\langle a_k, \widetilde{q}_j \rangle}{\langle \widetilde{q}_j, \widetilde{q}_j \rangle} \quad (j = 1, \ldots, k-1)$,
4. $\widetilde{q_k} := a_k - \sum\limits_{j=1}^{k-1} \widetilde{r_{jk}} \cdot \widetilde{q}_j$,
5. $\widetilde{r_{kk}} := 1 \quad$ (nem normálunk),

Az így nyert $\widetilde{q_1}, \ldots, \widetilde{q_n} \in \mathbb{R}^n$ vektorrendszer ortogonális.

Megj.: Levezetése teljesen hasonló. Kézi számolásra alkalmasabb. Ne felejtsünk el normálni...

Normálás utólag:

- $A = \widetilde{Q}\widetilde{R}$,
- $D := \widetilde{Q}^\top \widetilde{Q}$, azaz $D = \operatorname{diag}\left(\langle q_1, q_1 \rangle, \ldots \langle q_n, q_n \rangle\right)$,
- $A = \underbrace{\widetilde{Q} \cdot \sqrt{D}^{-1}}_{Q} \cdot \underbrace{\sqrt{D} \cdot \widetilde{R}}_{R} = Q \cdot R$,

  azaz $\widetilde{Q}$ oszlopait, mint vektorokat leosztjuk azok hosszával (normáljuk őket), $\widetilde{R}$ sorait pedig szorozzuk ugyanezekkel az értékekkel.
- Közvetlenül a $\sqrt{D} = \operatorname{diag}\left(\|q_1\|_2, \ldots, \|q_n\|_2\right)$ alakkal is dolgozhatunk.

Tétel: A Gram–Schmidt-ortogonalizáció műveletigénye

A szorzások és osztások száma

$$2n^3 + \mathcal{O}(n^2),$$

valamint n darab négyzetgyökvonás is szükséges.

Biz.: A k-adik lépésben:

skaláris szorzatok (r_{jk})	$(k-1)(2n-1)$
ortogonális vektor (s_k)	$(k-1)n + (k-1)n = (k-1)2n$
hossz (r_{kk})	$2n-1$
osztás (q_k)	n

Összesen:
$(k-1)(4n-1) + 3n - 1 = 4kn - 4n - k + 1 + 3n - 1 = 4kn - n - k,$

$$\sum_{k=1}^{n}(4kn - n - k) = 4n\sum_{k=1}^{n}k - n^2 - \sum_{k=1}^{n}k =$$
$$= 4n \cdot \frac{n(n+1)}{2} - n^2 - \frac{n(n+1)}{2} = 2n^3 + \mathcal{O}(n^2).$$

$\square$

Példa: *QR*, Gram–Schmidt

Készítsük el a következő mátrix QR-felbontását Gram–Schmidt-ortogonalizációval.

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

Gram–Schmidt ortogonalizációval normálással:

$$A = \begin{bmatrix} a_1 & a_2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} q_1 & q_2 \end{bmatrix} \cdot \begin{bmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{bmatrix} = Q \cdot R.$$

1. lépés: $a_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$-ből meghatározzuk r_{11}, q_1-et:

$$r_{11} = \|a_1\|_2 = \sqrt{1^2 + 2^2} = \sqrt{5}$$

$$q_1 = \frac{1}{r_{11}} a_1 = \frac{1}{\sqrt{5}} a_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

2. lépés:

$$A = \begin{bmatrix} a_1 & a_2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} q_1 & q_2 \end{bmatrix} \cdot \begin{bmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{bmatrix} = Q \cdot R$$

$a_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$-ből meghatározzuk r_{12}, r_{22}, q_2-t:

$$r_{12} = \langle a_2, q_1 \rangle = \left\langle \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\rangle = \frac{1}{\sqrt{5}}(2 \cdot 1 + 1 \cdot 2) = \frac{4}{\sqrt{5}}$$

$$\begin{aligned} s_2 = a_2 - r_{12} q_1 &= \begin{bmatrix} 2 \\ 1 \end{bmatrix} - \frac{4}{\sqrt{5}} \cdot \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \\ &= \frac{1}{5} \begin{bmatrix} 10 - 4 \\ 5 - 8 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 6 \\ -3 \end{bmatrix} = \frac{3}{5} \begin{bmatrix} 2 \\ -1 \end{bmatrix} \end{aligned}$$

$$r_{22} = \|s_2\|_2 = \left\| \frac{3}{5} \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right\|_2 = \frac{3}{5} \cdot \sqrt{2^2 + (-1)^2} = \frac{3}{5} \cdot \sqrt{5} = \frac{3}{\sqrt{5}}$$

$$q_2 = \frac{1}{r_{22}} s_2 = \frac{3}{5} \cdot \frac{\sqrt{5}}{3} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

□

Tehát a Q és R mátrixok a következők:

$$Q = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}, \quad R = \begin{bmatrix} \sqrt{5} & \frac{4}{\sqrt{5}} \\ 0 & \frac{3}{\sqrt{5}} \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} 5 & 4 \\ 0 & 3 \end{bmatrix}$$

Tartalomjegyzék

Rövid kitérő: vektorok „hossza”

Definíció: vektorok „hossza”

Az $\mathbb{R}^n$-beli v vektorok hagyományos értelemben vett hosszát, avagy „kettes normáját” jelölje $\|.\|_2$.
A következőképpen számolható:

$$\|v\|_2 := \sqrt{\langle v, v \rangle} = \sqrt{v^\top v} = \left(\sum_{k=1}^{n} v_i^2 \right)^{\frac{1}{2}}$$

Definíció: Householder-mátrix

A $H = H(v) \in \mathbb{R}^{n\times n}$ mátrixot *Householder-mátrixnak* nevezzük, ha

$$H(v) = I - 2vv^\top,$$

ahol $v \in \mathbb{R}^n$ és $\|v\|_2 = 1$.

Megjegyzés:

- A $H(v)$ transzformációs mátrixot nem kell előállítani, enélkül alkalmazzuk vektorokra, ez a Householder-transzformáció:
- $x \in \mathbb{R}^n$-re $H(v)x = (I - 2vv^\top)x = x - 2v\underbrace{(v^\top x)}_{\in\mathbb{R}}$.
- $y \in \mathbb{R}^n$-re $\quad y^\top H(v) = y^\top(I - 2vv^\top) = y^\top - 2\underbrace{(y^\top v)}_{\in\mathbb{R}}v^\top$.
- Mindkét esetben $4n$ művelet kell a mátrixszal való szorzás $2n^2 + \mathcal{O}(n)$-es műveletigénye helyett.

Állítás: Householder-mátrixok tulajdonságai

1. $H^\top = H$ (szimmetrikus),
2. $H^2 = I$, azaz $H^{-1} = H$ (ortogonális),
3. $H(v) \cdot v = -v$,
4. $\forall y \perp v : \quad H(v) \cdot y = y$.

Biz.: Használjuk ki, hogy $v^\top v = 1$ és $v^\top y = 0$.

1. $(I - 2vv^\top)^\top = I^\top - 2(v^\top)^\top v^\top = I - 2vv^\top$,
2. $(I - 2vv^\top)(I - 2vv^\top) = I - 2vv^\top - 2vv^\top + 4v \underbrace{v^\top v} v^\top = I$,
3. $(I - 2vv^\top)v = v - 2v \underbrace{v^\top v} = v - 2v = -v$,
4. $(I - 2vv^\top)y = y - 2v \underbrace{v^\top y} = y$. $\square$

Megjegyzés:

- $H(v)$ tükröző mátrix, a v-re merőleges (azaz v normálvektorú) $n-1$ dimenziós altérre (0-n átmenő egyenesre, síkra stb.) tükröz.
- Legyen $v \in \mathbb{R}^n$ és $\|v\|_2 = 1$, tetszőleges $x \in \mathbb{R}^n$ vektort bontsunk v-re merőleges és v-vel párhuzamos komponensekre: $x = a + b$, ahol $a \perp v$ és $b || v$. Ekkor az előző tétel utolsó két állítása alapján

$$H(v)x = H(v)a + H(v)b = a - b.$$

- Mivel $H(v)$ ortogonális mátrix, $\|H(v)x\|_2 = \|x\|_2$, vagyis a transzformáció a vektor hosszát nem változtatja meg.

Tétel: tetszőleges tükrözés Householder-mátrixszal

Legyen $a, b \in \mathbb{R}^n$, $a \neq b$ és $\|a\|_2 = \|b\|_2 \neq 0$. Ekkor a

$$v = \pm \frac{a-b}{\|a-b\|_2} \text{ választással } H(v) \cdot a = b.$$

Biz.: Ismerve, hogy $H(v) = I - 2vv^\top$, számoljuk végig a $H(v) \cdot a$ szorzatot. Közben használjuk ki, hogy $\|a\|_2 = \|b\|_2$, azaz $a^\top a = b^\top b$, valamint a skaláris szorzás kommutatív, azaz $a^\top b = b^\top a$.

$$\left(I - 2\frac{(a-b)(a-b)^\top}{\|a-b\|_2^2}\right) \cdot a = a - \frac{2(a-b)(a^\top a - b^\top a)}{(a-b)^\top (a-b)} =$$
$$= a - \frac{2(a-b)(a^\top a - b^\top a)}{a^\top a - a^\top b - b^\top a + b^\top b} = a - \frac{2(a-b)(a^\top a - b^\top a)}{2(a^\top a - b^\top a)} =$$
$$= a - (a-b) = b.$$

Tehát valóban, két különböző, de azonos hosszúságú vektor átvihető egymásba egy Householder-transzformáció által. □

Megjegyzés: Egyébként $H(v) \cdot b = a$ is teljesül.

Példa: Householder-féle tükrözés

Határozzuk meg azt a Householder-féle transzformációt, amely az azonos hosszúságú a, b vektorhoz előállítja azt a v vektort, melyre $H(v) \cdot a = b$. Ellenőrzésképpen végezzük is el a transzformációt.

$$a = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}, \qquad b = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}.$$

$$a - b = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

$\|a - b\|_2 = \sqrt{1^2 + (-2)^2 + 1^2} = \sqrt{6}$

Tehát $v = \frac{a-b}{\|a-b\|_2} = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ jó választás.

Ellenőrizzük végezzük el a transzformációt a-n:
$H(v) \cdot a = a - 2v \underbrace{(v^\top a)}_{\in \mathbb{R}} = a - 2(v^\top a)v.$

$$H(v) \cdot a = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} - 2 \cdot \frac{1}{\sqrt{6}} \underbrace{\begin{bmatrix} 1 & -2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}}_{3} \cdot \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} - \frac{6}{6} \cdot \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = b \quad \checkmark$$

$\square$

Példa: Householder-féle tükrözés

Határozzuk meg azt a Householder-féle transzformációt, amely a következő a vektort $b = k \cdot e_1$ alakúra hozza. Ellenőrzésképpen végezzük is el a transzformációt.

$$a = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}$$

A jó előjel választás σ-nak -1, mert a első eleme pozitív.

$$\sigma = -\|a\|_2 = -\sqrt{2^2 + (-2)^2 + 1^2} = -3$$

Ezzel az előjel választással stabilabb lesz az osztásunk v előállításban.

$$a - \sigma e_1 = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} - (-3) \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}$$

Látjuk, hogy valójában egyetlen műveletet kellett elvégeznünk a vektor első elemén. Ezzel a σ előjelválasztással elérjük, hogy $\|a - \sigma e_1\|_2 \geq \|a\|_2$.

$$\|a - \sigma e_1\|_2 = \sqrt{5^2 + (-2)^2 + 1^2} = \sqrt{30}$$

$$v = \frac{a - \sigma e_1}{\|a - \sigma e_1\|_2} = \frac{1}{\sqrt{30}} \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} \quad \text{jó választás.}$$

Ellenőrizzük végezzük el a transzformációt a-n:
$H(v) \cdot a = a - 2v\underbrace{(v^\top a)}_{\in\mathbb{R}} = a - 2(v^\top a)v.$

$$H(v) \cdot a = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} - 2 \cdot \frac{1}{\sqrt{30}} \underbrace{\begin{bmatrix} 5 & -2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}}_{15} \cdot \frac{1}{\sqrt{30}} \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} - \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 \\ 0 \\ 0 \end{bmatrix} = \sigma \cdot e_1 \quad \checkmark$$

$\square$

Tartalomjegyzék

Felső háromszög alakra hozás

Módszer:

- Legyen adott az $A \in \mathbb{R}^{n\times n}$ invertálható mátrix, első oszlopát jelölje a_1.
- Egy lépésben egy oszlopot kinullázunk a főátló alatt. ($\sim$ GE)
- Így $n-1$ lépésben felső háromszög alakot nyerünk.

Definíció: előjel függvény

$$\mathrm{sgn}: \mathbb{R} \to \mathbb{R}, \qquad \mathrm{sgn}\,(x) = \begin{cases} 1 & \text{ha } x > 0 \\ 0 & \text{ha } x = 0 \\ -1 & \text{ha } x < 0 \end{cases}$$

Megjegyzés: most, a Householder-transzformációknál nem engedhetjük meg a 0 értéket, helyette akár $+1$-et, akár -1-et választhatunk.

1. lépés:

$a_1 \Rightarrow \sigma_1 \cdot e_1$, ahol $\sigma_1 := -\operatorname{sgn}(a_{11}) \cdot \|a_1\|_2$ (tehát $|\sigma_1| = \|a_1\|_2$),

$$v_1 := \frac{a_1 - \sigma_1 e_1}{\|a_1 - \sigma_1 e_1\|_2}, \qquad H_1 := H(v_1).$$

Ekkor

$$H_1 \cdot A = H(v_1) \cdot A = \begin{pmatrix} \sigma_1 & * & \dots & * \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix}.$$

Megjegyzés: σ_1 megválasztásáról… így stabilabb.

2. lépés:

$b_1 \Rightarrow \sigma_2 \cdot e_1$, ahol $\sigma_2 := -\operatorname{sgn}(b_{11}) \cdot \|b_1\|_2$ (tehát $|\sigma_2| = \|b_1\|_2$),

$$\widetilde{v_2} := \frac{b_1 - \sigma_2 e_1}{\|b_1 - \sigma_2 e_1\|_2} \in \mathbb{R}^{n-1}$$

Ekkor

$$H(\widetilde{v_2}) \cdot B = \begin{pmatrix} \sigma_2 & * & \cdots & * \\ 0 & & & \\ \vdots & & C & \\ 0 & & & \end{pmatrix} \in \mathbb{R}^{(n-1)\times(n-1)}$$

2. lépés (teljes méretben ($n \times n$) felírva):

$$v_2 := \begin{pmatrix} 0 \\ \widetilde{v_2} \end{pmatrix}, \quad H_2 := H(v_2) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & H(\widetilde{v_2}) & \\ 0 & & & \end{pmatrix}.$$

Ekkor

$$H_2 \cdot H_1 \cdot A = \begin{pmatrix} \sigma_1 & * & * & \dots & * \\ 0 & \sigma_2 & * & \dots & * \\ 0 & 0 & & & \\ \vdots & \vdots & & C & \\ 0 & 0 & & & \end{pmatrix}.$$

Általában, k. lépés:
kinullázzuk az elemeket a főátló alatt a k. oszlopban.

Az ezt megvalósító transzformáció:

$$v_k := \begin{pmatrix} 0_{(1.)} \\ \vdots \\ 0_{(k-1).} \\ \widetilde{v_k} \end{pmatrix} \in \mathbb{R}^n, \quad H_k := H(v_k) = \begin{pmatrix} I_{k-1} & 0 \\ 0 & H(\widetilde{v_k}) \end{pmatrix}.$$

A gyakorlatban csak az $(n-k+1) \times (n-k+1)$-s mátrix részen dolgozunk a k. lépésben, mint a GE-nál. Az $(n-1)$-edik lépés után felső háromszög alakot kapunk.

A Householder-transzformáció alkalmazásai

Egyetlen LER megoldása:

$$Ax = b$$
$$H_1 \cdot A \cdot x = H_1 \cdot b$$
$$\vdots$$
$$\underbrace{H_{n-1} \cdots H_1 \cdot A}_{R} \cdot x = \underbrace{H_{n-1} \cdots H_1 \cdot b}_{d}$$
$$R \cdot x = d \to x \text{ (visszahelyettesítés)}$$

Ugyanúgy dolgozunk, mint a GE-nál. Végrehajtjuk a transzformációt az oszlopokon:

$$[A|b] \;\to \text{n-1 db H-trf.} \to\; [R|d] \;\to \text{visszahely.}$$

Mindig egyre kisebb méretű mátrixon dolgozunk a transzformációk során.

A Householder-transzformáció alkalmazásai

QR-felbontás készítése:

$$\underbrace{H_{n-1}\cdots H_2\cdot H_1}_{Q^{-1}=Q^\top}\cdot A = R$$

$$A = \underbrace{H_1\cdot H_2\cdots H_{n-1}}_{Q}\cdot R = Q\cdot R$$

Megfigyelhetjük, hogy Q előállításakor mindig a jobb oldalról végezzük a transzformációt, ekkor sorokra alkalmazzuk.

Az algoritmus: Q előállítására

$$\begin{aligned} & Q_0 = I \\ k=1,\ldots,n-1: \quad & Q_k := Q_{k-1}H_k \\ & Q := Q_{n-1} \end{aligned}$$

Tétel: QR-felbontás Householder-módszerrel

Invertálható mátrixok QR-felbontása elkészíthető $n-1$ db Householder-transzformáció segítségével.

Biz.: Láttuk. $\square$

Összefoglalva: A k. lépésben kinullázzuk a k. oszlop főátló alatti elemeit egy H_k ortogonális transzformáció segítségével, melyet a mátrix oszlopaira alkalmazunk a jobb alsó
$(n-k+1)\times(n-k+1)$-s mátrix részen.
A Q mátrixot úgy kapjuk, hogy egy egységmátrixból indulva a k. lépésben a H_k transzformációt jobbról alkalmazzuk a sorokra csak a jobb alsó $(n-k+1)\times(n-k+1)$-s mátrix részen.
$n-1$ lépés után megkapjuk felső háromszög alakot (R) és Q-t.

Tétel: A Householder-trf. műveletigénye LER-re

A LER megoldásának műveletigénye Householder-transzformációkkal:

$$\frac{4}{3}n^3 + \mathcal{O}(n^2),$$

valamint $2(n-1)$ darab négyzetgyökvonásra is szükség van.

Biz.:
A k-adik lépésben ($n - k + 1 =: h_k$ hosszú vektorokkal dolgozunk):

hossz (σ)	$2h_k - 1$,
normálvektor ($a - \sigma e_1, \|\|.\|\|_2, v$)	$1 + (2h_k - 1) + h_k = 3h_k$,
transzformáció ($(h_k - 1) + 1$ vektorra)	$h_k \cdot 4h_k$.

Összesen: $4h_k^2 + 5h_k - 1$, $(n - k =: s)$

$$\sum_{k=1}^{n-1} \left(4h_k^2 + 5h_k - 1\right) = \sum_{s=2}^{n} 4s^2 + \sum_{s=2}^{n} s + (n-1) = \frac{4}{3}n^3 + \mathcal{O}(n^2).$$

A visszahelyettesítés műveletigénye $n^2 + \mathcal{O}(n)$, belefér az előző alakba. $\square$

Tétel: A Householder-trf. műveletigénye QR-felbontásra

A QR-felbontás előállításának műveletigénye Householder-transzformációkkal:

$$\frac{8}{3}n^3 + \mathcal{O}(n^2),$$

valamint $2(n-1)$ darab négyzetgyökvonásra is szükség van.

Biz.:
A k-adik lépésben ($n-k+1 =: h_k$ hosszú vektorokkal dolgozunk):

hossz (σ)	$2h_k - 1$,
normálvektor ($a - \sigma e_1, \|.\|_2, v$)	$1 + (2h_k - 1) + h_k = 3h_k$,
transzformáció ($(h_k - 1) + h_k$ vektorra)	$(2h_k - 1) \cdot 4h_k$.

Összesen: $(8h_k^2 - 4h_k) + (5h_k - 1) = 8h_k^2 + h_k - 1$, $(n-k =: s)$

$$\sum_{k=1}^{n-1} \left(8h_k^2 + h_k - 1\right) = \sum_{s=2}^{n} 8s^2 + \sum_{s=2}^{n} s - (n-1) = \frac{8}{3}n^3 + \mathcal{O}(n^2).$$

$\square$

Megjegyzés: Ez kicsit több, mint a Gram–Schmidt-féle ortogonalizációnál, viszont ez a módszer numerikusan stabilabb.

Példák Matlab-ban

1. A Gram–Schmidt-féle ortogonalizációs eljárás működésének szemléltetése $\mathbb{R}^3$-beli vektorrendszer esetén.
2. Példák Householder-mátrixokra ($n \approx 3, 10, 20, 50$).
3. Példák Householder-transzformációra.
4. QR-felbontás készítése Householder módszerével ($n \approx 3, 7, 50, 100$).