QR rozklad (faktorizace)

- základní myšlenka: sloupce matice A jsou vektory, jejichž lineární obal tvoří obor hodnot $R(A)$ zobrazení Ax, když volíme x libovolně

neboli

$$\left(a_1 \middle| a_2 \middle| \cdots \middle| a_n\right)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 \left(a_1\right) + x_2 \left(a_2\right) + \dots + x_n \left(a_n\right)$$

A je matice $m \times n$, $m \geq n$

- vektory $a_1, \dots, a_n$ dávají postupně podprostory

$$\mathcal{L}(a_1),\ \mathcal{L}(a_1, a_2),\ \dots\ \mathcal{L}(a_1, \dots, a_n) = R(A)$$

lineární obal

přičemž v nich můžeme postupně konstruovat ortonormální bázi $q_1, \dots, q_n$ (za předpokladu úplné hodnosti matice A) takovou, že $\mathcal{L}(a_1) = \mathcal{L}(q_1)$, $\mathcal{L}(a_1, a_2) = \mathcal{L}(q_1, q_2)$ atd.

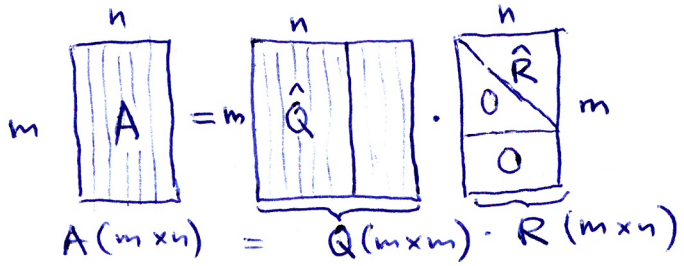
přičemž můžeme a_i vyjádřit např. takto (v bázi q_i)

$$\begin{aligned} a_1 &= r_{11} q_1 \\ a_2 &= r_{12} q_1 + r_{22} q_2 \\ &\vdots \\ a_n &= r_{1n} q_1 + \dots + r_{nn} q_n \end{aligned} \iff \left(a_1 \middle| \cdots \middle| a_n\right) = \left(q_1 \middle| \cdots \middle| q_n\right)\begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & & \ddots & \\ 0 & \cdots & 0 & r_{nn} \end{pmatrix}$$

tj. $A = \hat{Q} \cdot \hat{R}$, kde $\hat{Q}$ je obecně část (pro $m > n$) ortonormální matice Q splňující $Q^T Q = Q Q^T = \mathbb{1}$ a $\hat{R}$ je horní trojúhelníková matice

- tomuto rozkladu se říká <u>redukovaný QR rozklad</u> (reduced)

a pokud $\hat{Q}$ doplníme na Q lib. vektory ortogonálními na $q_1, \dots, q_n$ a $\hat{R}$ doplníme řádky s nulami, dostaneme <u>plný QR rozklad</u>



- lze ukázat, že každá matice $A \in \mathbb{C}^{m \times n}$ $(m \geq n)$ má QR rozklad, a tedy i $\hat{Q}\hat{R}$
  (přímou konstrukcí pomocí Gramovy-Schmidtovy ortogonalizace)

  a pokud je hodnost $A = n$ (maximální), pak existuje právě jeden rozklad $\hat{Q} \cdot \hat{R}$ s $r_{jj} > 0$, tj. r_{jj} je reálné, kladné, čímž fixujeme „fázi" q_j

  - Pokud je hodnost $A < n$, pak pro určité j budou vektory $q_1, \dots, q_j$ lineárně závislé, pak bude $r_{jj} = 0$ a jako q_j volíme lib. normovaný vektor kolmý na $q_1, \dots q_{j-1}$

- <u>použití na řešení $Ax = b$</u>
  - máme-li matici $(m \times m)$ A a její rozklad $Q \cdot R$, pak řešení x dostaneme přímo zpětnou substitucí pro $Rx = Q^t b$, neboť Q^t je inverzní k Q

  [Pozn: u některých metod (např. Householderova triangularizace) se Q přímo nekonstruuje (z A dostáváme R) a v průběhu výpočtu R se počítá i $Q^t b$]

- <u>výpočet QR rozkladu</u>
  - několik metod

	počet operací
1) přímá Gramova-Schmidtova ortogonalizace - není obecně stabilní, ztrácí se ortogonalita kvůli zaokrouhlovacím chybám	$\sim 2mn^2$
2) modifikovaná G-S ortogonalizace - o něco lepší, ale také ne zcela přesné	$\sim 2mn^2$
3) <u>Householderova triangularizace</u> - nejlepší metoda, zpětně stabilní	$\sim 2mn^2 - \frac{2}{3}n^3$
4) Givensovy rotace - o něco pomalejší, ale má své uplatnění např. v iteračních metodách (viz např. metoda GMRES v letním semestru)	$\sim 3mn^2$

1) přímá konstrukce pomocí Gramovy-Schmidtovi ortogonalizace

- z $m \times n$ matice A dává redukovaný QR rozklad
- ortogonalizujeme postupně a_2 na q_1, pak a_3 na q_1 a q_2 atd.

  a zároveň normalizujeme (předpoklad plné hodnosti A, jinak nutno testovat, zda $r_{jj} = 0$ a případně volit q_j jinak)

pro $j = 1, \dots, n$

$q_j = a_j$

pro $i = 1, \dots, j-1$

$$\left.\begin{aligned} r_{ij} &= q_i^T a_j \\ q_j &= q_j - r_{ij} q_i \end{aligned}\right\}$$ ortogonalizace, pokud a_j blízko q_i, odečítání 2 blízkých čísel (vektorů) $\Rightarrow$ ztráta přesnosti a ortogonality

$$\left.\begin{aligned} r_{jj} &= \|q_j\|_2 \\ q_j &= q_j / r_{jj} \end{aligned}\right\}$$ normalizace

zde je $\sim 4m$ operací a opakuje se $\sim \frac{n^2}{2}$ krát

celkem $O(2mn^2)$

2) modifikovaná Gramova-Schmidtova metoda

- ve standardní G-S metodě jsou jednotlivé q_j dány $P_j a_j$, tj. projektorem $P_j = I - q_1 q_1^T - \dots - q_{j-1} q_{j-1}^T$ na podprostor kolmý na $\mathcal{L}(q_1, \dots, q_{j-1}) = \mathcal{L}(a_1, \dots, a_{j-1})$ (alespoň v přesné aritmetice)
- tento projektor lze napsat jako součin projektorů

$$P_j = P_{\perp q_{j-1}} P_{\perp q_{j-2}} \cdots P_{\perp q_2} P_{\perp q_1}, \quad \text{kde } P_{\perp q_i} = I - q_i q_i^T$$

díky ortogonalitě vektorů q_i

- takto je to numericky výhodnější, P_j má hodnost $m-(j-1)$, kdežto $P_{\perp q_i}$ mají stále hodnost $m-1$
- algoritmus modif. G-S metody:

pro $i = 1, \dots, n$

$q_i = a_i$ (inicializace všech q_i)

pro $i = 1, \dots, n$

$$\left.\begin{aligned} r_{ii} &= \|q_i\|_2 \\ q_i &= q_i / r_{ii} \end{aligned}\right\}$$ normalizace (opět neošetřujeme zda $r_{ii} = 0$)

pro $j = i+1, \dots, n$

$$\left.\begin{aligned} r_{ij} &= q_i^T q_j \\ q_j &= q_j - r_{ij} q_i \end{aligned}\right\}$$ postupná ortogonalizace všech vektorů q_j najednou (i zde ztráta přesnosti, ale méně)

opět $\sim 4m$ operací

i zde je celkem $O(2mn^2)$ operací

- pozn.: G-S ortogonalizace je vlastně <u>trojúhelníková ortogonalizace</u>, neboť jde přepsat jako $AR_1R_2 \dots R_n = \hat{Q}$

kde
$$R_1 = \begin{pmatrix} \frac{1}{r_{11}} & -\frac{r_{12}}{r_{11}} & \dots & -\frac{r_{1n}}{r_{11}} \\ 0 & 1 & 0 \dots & 0 \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} \quad \text{a} \quad R_i = \begin{pmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & 1 & & \\ & & \frac{1}{r_{ii}} & -\frac{r_{i,i+1}}{r_{ii}} \dots & -\frac{r_{in}}{r_{ii}} \\ 0 & & & 1 & 0 \\ & & & & \ddots \\ & & & & 1 \end{pmatrix}$$

a $\hat{R} = R_n^{-1} R_{n-1}^{-1} \dots R_1^{-1}$

myšlenka Householderovy a Givensovy metody je opačná, a sice jde o <u>ortogonální triangularizaci</u>

3) Householderova triangularizace

- nyní budeme převádět A na horní △ tvar pomocí ortogonálních matic $Q_1, \dots, Q_n$, tj.

$$\underbrace{Q_n \dots Q_1}_{Q^T} A = R$$

a dostaneme tak <u>plný QR rozklad</u>

- matice $Q_1, \dots, Q_n$ musí být takové, aby nulovaly prvky pod diagonálou, tj.

$$A \xrightarrow{Q_1} \underset{Q_1A}{\begin{pmatrix} \times & \times & \times & \dots & \times \\ 0 & \times & & & \vdots \\ \vdots & \times & & & \vdots \\ \vdots & \vdots & & & \vdots \\ 0 & \times & \dots & & \times \end{pmatrix}} \xrightarrow{Q_2} \underset{Q_2Q_1A}{\begin{pmatrix} \times & \times & \times & \dots & \times \\ 0 & \times & \times & & \vdots \\ 0 & 0 & \times & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \times & \dots & \times \end{pmatrix}} \quad \text{atd.} \quad \left(\begin{array}{c}\times \text{ zde značí} \\ \text{obecně nenulovou} \\ \text{hodnotu}\end{array}\right)$$

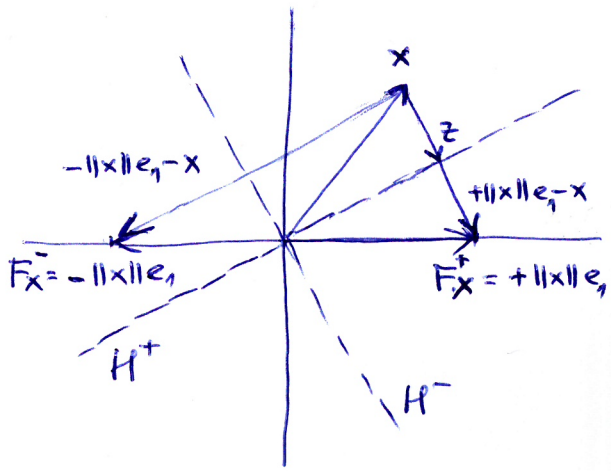
- to lze provést např. <u>Householderovým zrcadlením</u>

$$Q_k = \begin{pmatrix} I & 0 \\ 0 & F \end{pmatrix} \begin{matrix} \}\, k-1 \text{ řádků} \\ \}\, m-(k-1) \text{ řádků} \end{matrix}$$

kde F je takové, aby z vektoru

prvky od diagonály dolů v k-tém sloupci $\left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \right.$ $\xrightarrow{F}$ $\begin{pmatrix} \|x\|_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \|x\|_2 e_1$

jednotkový vektor



tj. F musí provést zrcadlení v „nadrovině" $H^+ \perp v = \|x\| e_1 - x$

(případně v „nadrovině $H^- \perp v = -\|x\| e_1 - x$, pokud je x příliš blízko $\|x\| e_1$)

- projekce do H^+ je dána projektorem $Px = \left(I - \frac{vv^t}{v^t v}\right) x = x - z$

ovšem my potřebujeme, aby $Fx = x - 2z$

neboli
$$Fx = \left(I - 2\frac{vv^t}{v^t v}\right) x$$

- vektor v se obvykle volí podle znaménka x_1,

tj. $v = -\mathrm{sign}(x_1)\|x\| e_1 - x$, ale protože F na znaménku vektoru v nezávisí, lze vzít jednoduše

$$v = \mathrm{sign}(x_1)\|x\| e_1 + x \qquad (\text{a např. } \mathrm{sign}(0) = 1)$$

- algoritmus pak upravuje A na R takto

pro $k = 1, \dots, n$

$x = A_{k:m,k}$ (proměnlivá délka x!)

$v_k = \mathrm{sign}(x_1)\|x\|_2 e_1 + x$ (volba v)

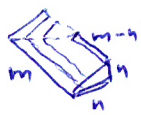
$v_k = v_k / \|v_k\|_2$ (normalizace v)

nejnáročnější část, úprava $\sim (m-k)\cdot(n-k)$ prvků, vždy 4 operace $\rightarrow$

$$A_{k:m,k:n} = A_{k:m,k:n} - 2 v_k (v_k^t A_{k:m,k:n})$$

dvojku lze „schovat" do v_k

celkem



objem $\frac{1}{2}mn^2 - \frac{1}{6}n^3$ krát 4 operace

$\sim 2mn^2 - \frac{2}{3}n^3$

matice Q se často nekonstruuje, ale počítá se buď ($Q^t = Q_n \dots Q_1$)

$Q^t b$: pro $k = 1, \dots n$

$b_{k:m} = b_{k:m} - 2v_k (v_k^t b_{k:m})$

nebo ($Q = Q_1 \dots Q_n$, neboť $Q_i^t = Q_i$)

Qb: pro $k = n, \dots, 1$

$b_{k:m} = b_{k:m} - 2v_k (v_k^t b_{k:m})$

(zde je nutné uchovat v_k!)

toto už je jen $\sim 4mn$ operací

pokud potřebujeme Q, působíme takto na jednotkovou matici (vektory e_j)

4) Givensovy rotace používají místo zrcadlení otočení, ale nebudeme je zde podrobněji probírat

- lze ukázat, že <u>Householderova triangularizace je zpětně stabilní</u>, tj. pro spočtené $\tilde{Q}$ a $\tilde{R}$ platí

$$\tilde{Q}\cdot\tilde{R} = A + \delta A \quad , \quad \frac{\|\delta A\|}{\|A\|} = \sigma(\varepsilon_M) \quad \text{pro vhodné } \delta A \in \mathbb{C}^{m\times n}$$

a též řešení soustavy $Ax = b$ je stabilní, pokud ji řešíme pomocí $Q\cdot R = A$ a následně $y = Q^{+}b$ a $x = R^{-1}y$.

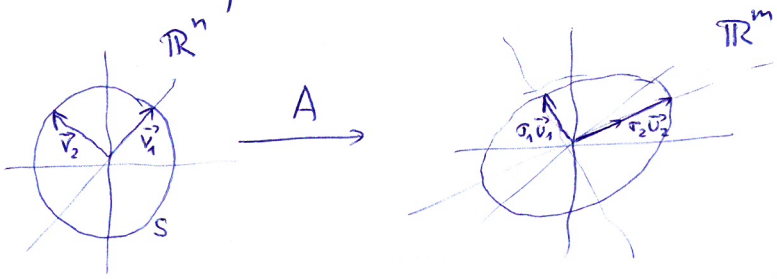
- ovšem pozor na podmíněnost úlohy, obecně bude jak $\tilde{Q}$ a $\tilde{R}$, tak $\tilde{x}$ mít chybu $\sigma(\varkappa(A)\varepsilon_M)$,

kde $\varkappa(A) = \|A\|\|A^{-1}\| = \frac{\sigma_1}{\sigma_m}$, kde σ_1 a σ_m jsou největší a nejmenší singulární hodnoty.

Singulární rozklad matice (Singular Value Decomposition – SVD)

- matice $m \times n$ A zobrazuje vektory x z $\mathbb{R}^n$ do prostoru $\mathcal{R}(A) \subset \mathbb{R}^m$

přitom jednotkové vektory z $\mathbb{R}^n$ (1-sféra) zobrazí na hyperelipsu v $\mathbb{R}^m$



stín, že některé směry můžou být degenerované $\sigma_i = 0$ (pro matice, které nemají hodnost n)

- lze najít ortonormální bázi v $\mathbb{R}^n$ (vektory $\vec{v}_1, \vec{v}_2 \dots$) takovou, že

$A\vec{v}_i = \sigma_i \vec{u}_i$, kde $\vec{u}_i$ jsou jednotkové vektory ve směrech hlavních poloos hyperelipsy

a σ_i budou seřazeny, tj. $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$

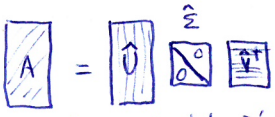
- maticově

$$\underset{m\times n}{A}\underset{n\times n}{\begin{pmatrix} v_1 & v_2 & \dots & v_n \end{pmatrix}} = \underset{m\times n}{\begin{pmatrix} u_1 & u_2 & \dots & u_n \end{pmatrix}} \begin{pmatrix} \sigma_1 & & 0 \\ & \sigma_2 & \\ 0 & & \ddots \; \sigma_n \end{pmatrix}$$

$\hat{V}$ je unitární
$\hat{U}$ splňuje $\hat{U}^\dagger \hat{U} = \mathbb{1}_{n\times n}$

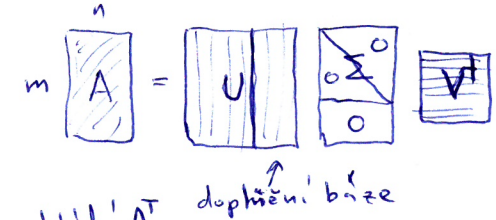
neboli

$$A\hat{V} = \hat{U}\hat{\Sigma} \quad \Rightarrow \quad A = \hat{U}\hat{\Sigma}\hat{V}^\dagger$$



σ_i jsou tzv. singulární hodnoty (reálné na rozdíl od vlastních čísel)
u_i jsou levé singulární vektory a v_i jsou pravé singulární vektory

- doplněním báze v $\mathbb{R}^m$ dostaneme tzv. plný sing. rozklad



tj. každá matice je diagonální ve vhodné bázi

- pro $m < n$ se rozkládá A^T

<u>Věta</u>: Každá matice $m \times n$ A má SVD rozklad. Navíc $\{\sigma_i\}$ jsou dány jednoznačně a pokud jsou navzájem různé, pak i $\{u_i\}$ a $\{v_i\}$ jsou dány jednoznačně až na komplexní fázi (± 1 u reálných matic)

<u>Dk</u>: Indukcí: nechť $\exists$ SVD pro matici $(m-1)\times(n-1)$. Položme $\sigma_1 = \|A\|_2 = \sup_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \sup_{\|x\|_2 = 1} \|Ax\|_2$

určitě existuje v_1 a u_1 takové, že

$\|v_1\|_2 = 1$, $\|u_1\|_2 = 1$ a $Av_1 = \sigma_1 u_1$

doplněním v_1 a u_1 na ortonormální bázi (libovolně) dostaneme U_1 a V_1 unitární matice, pro které

$$U_1^\dagger A V_1 = S = \begin{pmatrix} \sigma_1 & w^\dagger \\ 0 & B \end{pmatrix}$$

a platí $\left\| S \begin{pmatrix} \sigma_1 \\ w \end{pmatrix} \right\|_2 \geq \sigma_1^2 + w^\dagger w = (\sigma_1^2 + w^\dagger w)^{1/2} \left\| \begin{pmatrix} \sigma_1 \\ w \end{pmatrix} \right\|_2 \Rightarrow \|S\|_2 \geq (\sigma_1^2 + w^\dagger w)^{1/2}$

ale U_1, V_1 jsou unitární $\Rightarrow$

$\|S\|_2 = \|A\|_2 = \sigma_1 \rightarrow w = 0$, zbytek indukcí

Základní vlastnosti SVD rozkladu

- SVD vyjadřuje možnou změnu báze v $\mathbb{R}^m$ a $\mathbb{R}^n$ tak, aby

  $b = Ax$ přešlo na $b' = \Sigma x'$, kde $x' = V^\dagger x$ a $b' = U^\dagger b$

- rozdíl od vlastních hodnot a vektorů
  - ty mají smysl jen pro čtvercové matice a rozklad $A = X \Lambda X^{-1}$ existuje jen pro nedefektní matice (viz později), oba prostory ve stejné bázi
  - pokud je matice **hermitovská** $(A = A^\dagger)$, pak λ_i lze volit tak, že $\sigma_i = |\lambda_i|$

- hodnost matice A = hodnosti Σ (U i V mají plnou hodnost) ⇒

  ⇒ # nenulových σ_i určuje hodnost A

  pokud je r nenul. σ_i, pak $\mathcal{R}(A) = \mathcal{L}(u_1, \dots u_r)$, $\mathrm{Ker}(A) = \mathrm{null}(A) = \mathcal{L}(v_{r+1}, \dots, v_n)$

- $\|A\|_2 = \sigma_1 = \max \sigma_i = \|\Sigma\|_2$ a dále $\|A\|_F = \left(\sum_{i,j} |a_{ij}|^2 \right)^{1/2} = \sqrt{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_r^2}$

- sestrojme matici $n \times n$ $A^\dagger A$, pak

  $$A^\dagger A = (U \Sigma V^\dagger)^\dagger (U \Sigma V^\dagger) = V (\Sigma^\dagger \Sigma) V^\dagger \quad \text{a tedy}$$

  $A^\dagger A$ a $\Sigma^\dagger \Sigma$ mají stejná vlastní čísla $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$

- pro $m \times m$ matici A je $\det(A) = \prod_{i=1}^{m} \sigma_i$

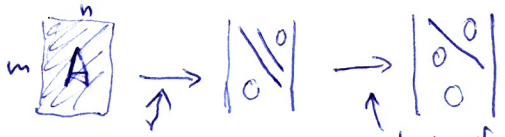
- rozepsáním Σ na součet matic $\begin{pmatrix} 0 & & \\ & \sigma_i & \\ & & 0 \end{pmatrix}$ dostaneme $A = \sum_{j=1}^{r} \sigma_j u_j v_j^\dagger$ ← počet nenulových σ_i

  ⇒ aproximace A omezením na velká σ_i

  – „nejlepší aproximace A" pomocí matice určité hodnosti

Výpočet SVD rozkladu (nástin)

- přímočaré by bylo najít vlastní čísla $A^\dagger A = V \Lambda V^\dagger$ a spočítat Σ jako odmocninu Λ, pak vyřešit $U\Sigma = AV$ (např. QR faktorizací)
  - ovšem ukazuje se, že jde o nestabilní alg., pokud některá $\sigma_k \ll \|A\|$
- stabilní algoritmus založen na převedení na bidiagonální matici

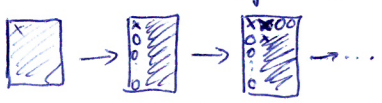


Golubova-Kahanova bidiagonalizace $O(mn^2)$

přesněji $4mn^2 - \frac{4}{3}n^3$

velmi efektivní iterační metody $O(n^2)$, založené na QR fakt. nebo rozděl a panuj metodě

např. Householderovy reflexe zleva i zprava



jde to i lépe pro $m \gg n$ až $2mn^2 + 2n^3$