Kap. 5. QR-Zerlegung und überbestimmte Gleichungssysteme

(19.4.2010)

In diesem Kapitel wird nur der $\mathbb{K}^n$ mit Standard-IP betrachtet.

5.1 QR-Zerlegung

Eine Umschreibung des Gram-Schmidtschen Orthogonalisierungsverfahrens in Matrizenschreibweise.

Seien $a_1, \dots, a_n \in \mathbb{K}^m$ l.u. $\quad (\Rightarrow n \leq m)$

Gram-Schmidt:

$$w_k = a_k - \sum_{j=1}^{k-1} \tilde{r}_{jk} \, w_j \qquad \text{mit} \quad \tilde{r}_{jk} = \frac{\langle w_j, a_k \rangle}{\langle w_j, w_j \rangle} \qquad (j<k)$$

$$1 \leq k \leq n.$$

$$\Longleftrightarrow \quad a_k = \sum_{j=1}^{k-1} \tilde{r}_{jk} \, w_j + 1 \cdot w_k + 0 \cdot w_{k+1} + \dots + 0 \cdot w_n$$

$$= \sum_{j=1}^{n} \tilde{r}_{jk} \, w_j =$$

$$\tilde{r}_{jk} = \begin{cases} \frac{\langle w_j, a_k \rangle}{\langle w_j, w_j \rangle} & \text{--- } j < k \\ 1 & \text{--- } j = k \\ 0 & \text{--- } j > k \end{cases}$$

$$= \sum_{j=1}^{n} w_j \, \tilde{r}_{jk}$$

$$(a_1 \dots a_n) = \underbrace{(w_1 \dots w_n)}_{\text{Orthogonalsystem}} \tilde{R} \qquad \text{mit} \quad \tilde{R} = (\tilde{r}_{jk})_{\substack{1 \le j \le n \\ 1 \le k \le n}}$$

$\tilde{R}$ ... obere Dreiecksmatrix.

Wir normieren das Orthogonalsystem $w_1, \dots, w_n$ zu einem Orthonormalsystem $e_1, \dots, e_n$

$$e_j = \frac{1}{\|w_j\|} w_j. \qquad w_j = \|w_j\| \cdot e_j$$

$$(a_1 \dots a_n) = (\|w_1\| \cdot e_1 \quad \dots \quad \|w_n\| \cdot e_n) \tilde{R} =$$

$$= \underbrace{(e_1 \; \dots \; e_n)}_{\substack{=: \tilde{Q} \\ \text{Orthonormalsystem}}} \cdot \underbrace{\mathrm{diag}(\|w_1\|, \dots, \|w_n\|) \cdot \tilde{R}}_{=: \tilde{\tilde{R}}}$$

j-te Zeile von $\tilde{\tilde{R}}$ ist $\|w_j\|$ mal j-te Zeile von $\tilde{R}$.

$\tilde{\tilde{R}}$ ist nach wie vor obere Dreiecksmatrix,

Diagonalelemente von $\tilde{\tilde{R}}$: $\|w_1\|, \dots, \|w_n\|$.

Ergänze ONS $(e_1, \dots, e_n)$ zu ONB $(e_1, \dots, e_n, e_{n+1}, \dots, e_m)$

Setze $Q := (e_1, \dots, e_n, e_{n+1}, \dots, e_m) \in K^{m \times m}$

Da Spalten von Q eine Orthonormalbasis sind, ist Q eine unitäre Matrix. (Proposition in Abschnitt 4.2).

$$\Rightarrow (a_1, \dots, a_n) = (e_1 \; \dots \; e_n) \cdot \tilde{\tilde{R}} =$$

$$= \underbrace{\begin{pmatrix} e_1 \; \dots \; e_n \; e_{n+1} \; \dots \; e_m \end{pmatrix}}_{=: Q} \cdot \underbrace{\begin{pmatrix} \tilde{\tilde{R}} \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{=: R}$$

Satz 5.1 (Reduzierte QR-Zerlegung) Sei $A \in \mathbb{K}^{m \times n}$ mit $\operatorname{rang} A = n$. Dann gibt es eine Matrix $\hat{\tilde{Q}} \in \mathbb{K}^{m \times n}$ mit $\hat{\tilde{Q}}^* \hat{\tilde{Q}} = I$ und eine obere Dreiecksmatrix $\hat{\tilde{R}} \in \mathbb{K}^{n \times n}$ mit

$$A = \hat{\tilde{Q}} \hat{\tilde{R}}$$

Die Diagonalelemente $\tilde{\tilde{r}}_{jj}$ von $\hat{\tilde{R}}$ geben den Abstand von

a_j vom $\operatorname{span}\{a_1, \dots, a_{j-1}\}$ an.

Beweis. Zerlegung oben beweisen

$$\tilde{r}_{jj} = \|w_j\| = \|a_j - P_{\operatorname{span}\{a_1, \dots, a_{j-1}\}}(a_j)\|$$

Normalabstand = Kürzester Abstand $\square$

Satz 5.2 (volle QR-Zerlegung) Sei $A \in \mathbb{K}^{m \times n}$ mit $\operatorname{rang} A = n$. Dann gibt es eine unitäre Matrix $Q \in \mathbb{K}^{m \times m}$ und eine obere Dreiecksmatrix $R \in \mathbb{K}^{m \times n}$, sodass

$$A = QR.$$

Beweis. Siehe oben. $\square$

Bemerkung.
- Gram-Schmidt liefert QR.
- Andere Verfahren (Householder, Givens) liefern auch QR-Zerlegung aber numerisch stabiler
- Aus QR Zerlegung erhalte Orthonormalbasis für Spaltenraum von A:
  Nehme erste n Spalten von Q.
- Aus $A \in \mathbb{K}^{m \times n}$ und $\operatorname{rang} A = n$ folgt automatisch $m \geq n$

22.4.2010

5.2 Überbestimmte Gleichungssysteme

Sei $A \in \mathbb{K}^{m \times n}$ und $b \in \mathbb{K}^m$ mit $\operatorname{rang} A = n$. Betrachte lin. GLS

$$Ax = b$$

Wenn b nicht besonders günstig $(\operatorname{rang}(A \,\vdots\, b) = \operatorname{rang} A = n)$ gibt es keine Lösung x.
Man könnte sich aber mit Näherungslösung zufrieden geben.

Definition Sei $A \in \mathbb{K}^{m \times n}$, $b \in \mathbb{K}^m$, $x \in \mathbb{K}^n$. Dann heißt x die Lösung des Gleichungssystems „$Ax=b$" im Sinne der kleinsten Fehlerquadrate („least squares solution"), wenn

$$\|Ax-b\|_2 = \min\{\|Ay-b\|_2 \mid y\in\mathbb{K}^n\}.$$

1. Zugang: Normalgleichungen.

Kürzester Abstand = Normalabstand
Wir wollen b durch ein Element von $\{Ay \mid y\in\mathbb{K}^n\} = \operatorname{Im} A = \text{Spaltenraum}(A)$ approximieren.

Somit: $\quad Ax = P_{\operatorname{Im} A}(b). \underset{\substack{\text{Korollar 6}\\ \text{zu Satz 4.5}}}{=} A(A^*A)^{-1}A^*\cancel{b}$

Wir möchten A von links „kürzen". (Achtung: A ist vermutlich nicht invertierbar. Wir brauchen aber nur, dass $x\mapsto Ax$ injektiv ist, d.h. $\operatorname{Ker} A=\{0\} \underset{\text{Dim. Form}}{\Longleftrightarrow} n=\dim \operatorname{Im} A \Longleftrightarrow \operatorname{rang} A = n$)

Wenn $\operatorname{rang} A=n$, folgt also

$$x = (A^*A)^{-1} A^* b$$

$\Longrightarrow$ $\underset{\substack{n \times n \\ \text{regulär}}}{(A^*A)}\, x = A^* b.$ $n \times n$ lin GLS.

$(A^*A)^{-1}A^*$ heißt „Pseudoinverse von A"

Satz 5.3. Seien $A \in \mathbb{K}^{m \times n}$, $b \in \mathbb{K}^m$, rang $A = n$. Dann gibt es eine eindeutige Lösung x des überbest. GLS „$Ax = b$" im Sinne der kleinsten Fehlerquadrate, nämlich die eindeutige Lösung der „Normalgleichungen"

$$(A^*A)x = A^*b$$

Bemerkung 1 „Multipliziere ‚$Ax = b$' von links mit A^* und erhalte die Normalgleichungen".

Beispiel aus Einführung. $(x_1, y_1), \ldots, (x_{200}, y_{200})$.

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + a_5 e^x + a_6 \cos x + a_7 \sin x$$

$a_0, \dots, a_7$ unbek.

j-te Gleichung $\quad f(x_j) = a_0 + a_1 x_j + \dots + a_7 \sin x_j = y_j$

$$\begin{pmatrix} 1 & x_j & x_j^2 & x_j^3 & x_j^4 & e^{x_j} & \cos x_j & \sin x_j \end{pmatrix} \begin{pmatrix} a_0 \\ \vdots \\ a_7 \end{pmatrix} = y_j$$

also

$$\begin{pmatrix} 1 & x_1 & x_1^2 & x_1^3 & x_1^4 & e^{x_1} & \cos x_1 & \sin x_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{200} & x_{200}^2 & x_{200}^3 & x_{200}^4 & e^{x_{200}} & \cos x_{200} & \sin x_{200} \end{pmatrix} \begin{pmatrix} a_0 \\ \vdots \\ a_7 \end{pmatrix} \overset{!}{=} \begin{pmatrix} y_1 \\ \vdots \\ y_{200} \end{pmatrix}$$

<u>Bemerkung 2.</u> Schränkt man sich bei den Ansatzfunktionen auf Polynome ein

$$f(x) = a_0 + a_1 x + \dots + a_k x^k,$$

so hat die Systemmatrix A die Gestalt

$$A = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^k \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_N & x_N^2 & \cdots & x_N^k \end{pmatrix} = \left(x_i^{j-1} \right)_{\substack{1 \le i \le N \\ 1 \le j \le k+1}}.$$

$$A^t = \begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_N \\ & \vdots & \\ x_1^k & \cdots & x_N^k \end{pmatrix}$$

$$A^tA = \begin{pmatrix} N & \sum x_i & \sum x_i^2 & \cdots & \sum x_i^k \\ \sum x_i & \sum x_i^2 & \sum x_i^3 & \cdots & \sum x_i^{k+1} \end{pmatrix} \qquad A^t y = \begin{pmatrix} \sum y_i \\ \sum x_i y_i \end{pmatrix}$$

$$\begin{pmatrix} \vdots & & \\ \sum x_i^k & \sum x_i^{2k} \end{pmatrix} \qquad \begin{pmatrix} \vdots \\ \sum x_i^k y_i \end{pmatrix}$$

übliche Notation:

$$[1] = \sum 1 \qquad [y] = \sum y_i$$
$$[x] = \sum x_i \qquad [xy] = \sum x_i y_i$$
$$[xx] = \sum x_i^2$$

Normalgleichungen

$$\begin{pmatrix} [1] & [x] & [xx] & \dots & [x^k] \\ \vdots & & & & \\ [x^k] & & & & [x^{2k}] \end{pmatrix} z = \begin{pmatrix} [y] \\ [xy] \\ \vdots \\ [x^k y] \end{pmatrix}$$

Wenn man sich auf Geraden einschränkt „Regressionsgerade"

$$\begin{pmatrix} [1] & [x] \\ [x] & [xx] \end{pmatrix} \cdot z = \begin{pmatrix} [y] \\ [xy] \end{pmatrix}$$

Cramersche Regel

$$a_0 = \frac{\begin{vmatrix} [y] & [x] \\ [xy] & [xx] \end{vmatrix}}{} = \frac{[y][xx] - [xy][x]}{}$$

$$\begin{vmatrix} [1] & [x] \\ [x] & [xx] \end{vmatrix} \qquad [1][xx] - [x]^2$$

$$a_1 = \ldots$$

Bemerkung 3. Normalgleichungen können numerisch problematisch sein, meist schlecht konditioniert.

Variante 2. Überbest. GLS via QR

$A \in \mathbb{K}^{m \times n}$ $\quad b \in \mathbb{K}^m$ $\quad x$ gesucht, sodass $\|Ax - b\|_2$ minimal wird,

Annahme: $\operatorname{rang} A = n$.

QR-Zerlegung von A $\quad A = Q \cdot R$. $\quad Q$ unitär.

$$\|Ax - b\|_2^2 = \|QRx - \underbrace{QQ^*}_{I} b\|_2^2 = \|Q(Rx - Q^*b)\|_2^2$$

$$\underset{\substack{Q \text{ ist unitär} \\ \text{und daher längentreu}}}{=} \|Rx - Q^*b\|_2^2$$

$$R = \begin{array}{c} \left[\begin{array}{c} R_1 \\ \hline 0 \end{array}\right] \begin{array}{l} \}n \\ \\ \end{array} \\ \underbrace{\qquad}_{n} \end{array} \qquad Q^* b =: \left[\begin{array}{c} c \\ \hline d \end{array}\right] \begin{array}{l} \}n \\ \}m-n \end{array}$$

R_1 regulär

$$Rx = \left[\begin{array}{c} R_1 x \\ \hline 0x \end{array}\right] = \left[\begin{array}{c} R_1 x \\ \hline 0 \end{array}\right]$$

$$\|Ax - b\|_2^2 = \left\| \begin{pmatrix} R_1 x - c \\ -d \end{pmatrix} \right\|_2^2 = \|R_1 x - c\|_2^2 + \|-d\|_2^2$$

$\|Ax-b\|_2^2 \to$ min

$R_1 x = c$: $\|R_1 x - c\|_2^2 \to 0$

$\|-d\|_2^2$ konstant $\to \|-d\|_2^2$

Satz 5.4 Sei $A \in K^{m \times n}$, $b \in K^m$, rang $A = n$. Sei $A = Q \cdot R$ eine volle QR-Zerlegung von A, R_1 die ersten n Zeilen von R, c die ersten n Zeilen von $Q^* b$. Dann ist die eindeutige Lösung x

des überbest. GLS „$Ax=b$" (SdkFQ) die eindeutige Lösung des $n\times n$-Gleichungssystems

$$R_1 x = c$$

obere Dreiecksmatrix $\Rightarrow$ nur mehr Rückwärtseinsetzen

5.3. Unterbestimmte GLS

$A \in \mathbb{K}^{m\times n}$ $\quad b \in \mathbb{K}^m$ $\quad \operatorname{rang} A = m$, also $n \geq m$

Jetzt hat $Ax=b$ unendlich viele Lösungen.
Suche jene Lösung x von $Ax=b$, die $\|x\|_2$ minimiert.

Bestimme QR-Zerlegung von A^*, also $A^* = QR$, also $A = R^* Q^*$

$$Ax=b \iff R^*\underbrace{(Q^*x)}_{=:y} = b \iff R^* y = b \text{ und } x = Qy$$

Es gilt $\|x\|_2 = \|Qy\|_2 = \|y\|_2$. Wir minimieren $\|y\|_2$.

Q längentreu

$$R = \begin{pmatrix} R_1 \\ \hdashline 0 \end{pmatrix} \begin{array}{l} \}\, m \\ \}\, n-m \end{array} \qquad\qquad y = \begin{pmatrix} y_1 \\ \hdashline y_2 \end{pmatrix} \begin{array}{l} m \\ \}\, n-m \end{array}$$

(the columns of R: m)

$$b = R^* y = \left(R_1^* \;\vdots\; 0 \right) \begin{pmatrix} y_1 \\ \hdashline y_2 \end{pmatrix} = R_1^* y_1 + 0 y_2 = \underbrace{R_1^*}_{\text{regulär}} y_1 ,$$

y_1 eindeutig durch $m \times m$ - GLS $\quad R_1^* y_1 = b \quad$ bestimmt,

y_2 für Gleichung egal, wähle zur Minimierung von $\|y\|_2^2 = \|y_1\|_2^2 + \|y_2\|_2^2$

$$y_2 = 0.$$

$$x = Q \begin{pmatrix} (R_1^*)^{-1} b \\ 0 \end{pmatrix}$$

<u>Bemerkungen</u> 1) reguläres quadr. GLS $\quad Ax = b \quad$ auch via QR lösbar

$$QRx = b \qquad\qquad A = QR$$

$$\underline{Rx = Q^* b}$$

Dreiecksmatrix. (kann numerisch interessant sein)

2) Rang-Einschränkungen sind wesentlich
Ausweg: siehe letztes Kapitel der VO.

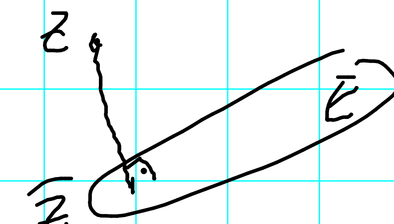
5.4. Noch 2 geometrische Bemerkungen im Zusammenhang mit Orthogonalität

1. Kürzeste Abstände zu affinen Hyperebenen

$\mathbb{R}^n$; affine Hyperebene = affiner Teilraum der Dimension $n-1$, also Ebenen im $\mathbb{R}^3$ und Geraden im $\mathbb{R}^2$.

$$E = \{ x \in \mathbb{R}^n \mid \langle n, x \rangle = \langle n, c \rangle \},$$ wobei n normierter Normalvektor und $c \in E$.

$z \in \mathbb{R}^n$. Gesucht: kürzester Abstand von z ~~zu~~ E.



$z - \bar{z} \parallel n$

Verschiebe alles um $-c$: $\quad z-c$ und $\underbrace{E-c}_{\text{Untervektorraum}}$,

Wir müssen also $z-c$ ~~auf~~ zerlegen in $(E-c) \oplus \underbrace{(E-c)^{\perp}}_{\operatorname{span}\{n\}}$

$$\text{Normalabstand} = \left\| P_{\operatorname{span}\{n\}}(z-c) \right\|_2 = \left\| \langle n, z-c \rangle n \right\|_2 =$$

$$= |\langle n, z-c \rangle| \cdot \underbrace{\|n\|_2}_{1} =$$

$$= |\langle n, z-c \rangle|$$

diese Abstandsformel findet man immer wieder.

2. Norm des Kreuzprodukts

33

Es gab ein Übungsbeispiel, dass

$\|a \times b\| = \|a\| \, \|b\| \sin \sphericalangle(a,b)$, = Fläche des von a und b aufgesp. Parallelograms

Dort aus Lagrangescher Identität hergeleitet.

Alternative:

$$\|a\times b\|_2^2 = \langle \underbrace{a\times b, a\times b}_{\text{Spatprod}} \rangle = \left|\det\left(a\times b, a, b\right)\right| =$$

$$= \text{Volumen}\left(a\times b, a, b\right) \overset{\uparrow}{=} \underbrace{\|a\times b\|_2}_{\text{Höhe}} \cdot \underbrace{\text{Fläche}(a, b)}_{\text{Grundfläche}}$$

weil $a\times b$ normal auf a und b steht

$$\Rightarrow \quad \|a\times b\|_2 = \text{Fläche}(a, b)$$

✓