Απλή Παραγοντοποίηση QR

ΔΠΜΣ Πληροφορικής και Υπολογιστικής Βιοϊατρικής

Δρ. Αρετάκη Αικατερίνη

Απρίλιος 2024

Βασικές Έννοιες

- *Βαθμός* (rank) ενός πίνακα $A \in \mathbb{R}^{m \times n}$ $\Rightarrow$ πλήθος γραμμικά ανεξάρτητων διανυσμάτων που ορίζονται από τις γραμμές ή τις στήλες του πίνακα A.
- *Νόρμα διανυσμάτων* $\Rightarrow$ Μια συνάρτηση $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ που για κάθε $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ ικανοποιεί τα ακόλουθα:
  1. $\|\mathbf{x}\| \geq 0$ (μη αρνητική).
  2. $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = 0$
  3. $\|\lambda\mathbf{x}\| = |\lambda|\|\mathbf{x}\|$ για κάθε $\lambda \in \mathbb{R}$
  4. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{x}\|$ (τριγωνική ανισότητα).
- *Ευκλείδια νόρμα* διανύσματος $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \Rightarrow$

$$\|\mathbf{x}\|_2 = \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2} = \mathbf{x}^T\mathbf{x}$$

Η Ευκλείδεια νόρμα επάγεται από το Ευκλείδειο εσωτερικό γινόμενο

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1y_1 + x_2y_2 + \cdots x_ny_n,$$

δηλαδή

$$\|\mathbf{x}\|_2 = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$$

Βασικές Έννοιες

- Αν $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^n$ είναι γραμμικά ανεξάρτητα διανύσματα και

$$\mathrm{proj}_{\mathbf{v}_1}(\mathbf{v}_2) = \frac{\langle \mathbf{v}_2, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1$$

  είναι η *προβολή* του $\mathbf{v}_2$ επί του $\mathbf{v}_1$, τότε

$$\left(\mathbf{v}_2 - \mathrm{proj}_{\mathbf{v}_1}(\mathbf{v}_2)\right) \perp \mathbf{v}_1 \Rightarrow \langle \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle} \mathbf{v}_1, \mathbf{v}_1 \rangle = 0$$

- Η μέθοδος Gram–Schmidt χρησιμοποιεί την παραπάνω ιδιότητα και μέσω γραμμικών συνδυασμών των προβολών μιας αρχικής βάσης κατασκευάζει ορθοκανονική βάση.

Ορθοκανονικοποίηση Gram–Schmidt

Έστω οτι $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$, $\mathbf{u}_j \in \mathbb{R}^m$ είναι μια βάση του $\mathbb{R}^n$. Τότε η κατασκευή πραγματοποιείται σε 2 στάδια:

1. Ορθογώνια βάση $\mathcal{O} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ (διανύσματα ανά δύο κάθετα):

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{u}_1 \\
\mathbf{v}_2 &= \mathbf{u}_2 - \mathrm{proj}_{\mathbf{v}_1}(\mathbf{u}_2) = \mathbf{u}_2 - \frac{\langle \mathbf{u}_2, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1 \\
\mathbf{v}_3 &= \mathbf{u}_3 - \mathrm{proj}_{\mathbf{v}_1}(\mathbf{u}_3) - \mathrm{proj}_{\mathbf{v}_2}(\mathbf{u}_3) = \mathbf{u}_3 - \frac{\langle \mathbf{u}_3, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1 - \frac{\langle \mathbf{u}_3, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle}\mathbf{v}_2 \\
\vdots &= \vdots \\
\mathbf{v}_n &= \mathbf{u}_n - \sum_{j=1}^{n-1} \mathrm{proj}_{\mathbf{v}_j}(\mathbf{u}_n) = \mathbf{u}_n - \sum_{j=1}^{n-1} \frac{\langle \mathbf{u}_n, \mathbf{v}_j \rangle}{\langle \mathbf{v}_j, \mathbf{v}_j \rangle}\mathbf{v}_j.
\end{aligned}$$

2. Κανονικοποιημένη βάση $\mathcal{U} = \{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ (μοναδιαία διανύσματα):

$$\mathbf{q}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}, \cdots, \mathbf{q}_n = \frac{\mathbf{v}_n}{\|\mathbf{v}_n\|}.$$

Απλή παραγοντοποίηση QR

Θεώρημα

Έστω πίνακας $A \in \mathbb{R}^{m\times n}$ $(m \geq n)$, *του οποίου οι στήλες* $\mathbf{a}_1, \mathbf{a}_2, \ldots,$ $\mathbf{a}_n \in \mathbb{R}^m$ *είναι γραμμικά ανεξάρτητα διανύσματα, δηλ.* $\operatorname{rank}(A) = n$.

Τότε υπάρχει πίνακας $Q \in \mathbb{R}^{m\times n}$ *με* $Q^TQ = I_n$ *και άνω τριγωνικός πίνακας* $R \in \mathbb{R}^{n\times n}$ *με θετικά διαγώνια στοιχεία ώστε*

$$A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & & r_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & r_{nn} \end{bmatrix} = QR.$$

- Κάθε $m \times n$ πίνακας A πλήρους βαθμού, δηλαδή $\operatorname{rank}(A) = n \leq m$, έχει μοναδική παραγοντοποίηση QR.

Παρατηρήσεις

- Οι στήλες $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n$ του $m \times n$ πίνακα Q είναι ορθοκανονικά διανύσματα του $\mathbb{R}^m$, τα οποία προκύπτουν άμεσα από την ορθοκανονικοποίηση Gram-Schmidt:

$$\text{span}\{\mathbf{q}_1, \ldots, \mathbf{q}_n\} = \text{span}\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}.$$

- Ο $n \times n$ πίνακας R έχει στοιχεία τους συντελεστές της ορθοκανονικοποίησης Gram-Schmidt:

$$\begin{aligned}
\mathbf{a}_1 &= r_{11}\mathbf{q}_1 \\
\mathbf{a}_2 &= r_{12}\mathbf{q}_1 + r_{22}\mathbf{q}_2 \\
&\vdots \\
\mathbf{a}_n &= r_{1n}\mathbf{q}_1 + r_{2n}\mathbf{q}_2 + \cdots + r_{nn}\mathbf{q}_n.
\end{aligned}$$

  Επιπλέον, είναι αντιστρέψιμος καθώς έχει θετικά διαγώνια στοιχεία ($r_{ii} > 0$, $i = 1, \ldots, n$).

Παρατηρήσεις

- Η διαδικασία ορθογωνιοποίησης Gram-Schmidt αποτελεί ένα πολύ καλό θεωρητικό εργαλείο, αλλά δεν είναι ένας καλός αριθμητικά αλγόριθμος όταν υλοποιείται στην κλασική του εκδοχή.

- Αν σε αριθμητική κινητής υποδιαστολής (floating-point) εφαρμοστεί η κλασική Gram-Schmidt μέθοδος σε ένα σύνολο διανυσμάτων, τα οποία δεν είναι ήδη κοντά στο να αποτελούν ένα ορθοκανονικό σύνολο, τότε το αποτέλεσμα μπορεί να είναι ένα νέο σύνολο διανυσμάτων τα οποία απέχουν πολύ από ένα ορθοκανονικό σύνολο.

Εφαρμογές της QR Παραγοντοποίησης

Η QR παραγοντοποίηση μπορεί να εφαρμοστεί στα ακόλουθα:

- Επίλυση γραμμικού συστήματος $Ax = b$:

  $$Ax = b \;\Leftrightarrow\; QRx = b \;\Leftrightarrow\; Qy = b \text{ και } y = Rx,$$

  όπου το σύστημα $y = Rx$ επιλύεται με προς τα πίσω αντικατάσταση και το σύστημα $Qy = b$ επιλύεται άμεσα θεωρώντας $y = Q^T b$.

- Υπολογισμός ορίζουσας και αντιστρόφου του πίνακα $A \in \mathbb{R}^{n \times n}$:

  $$\det A = \det Q \cdot \det R = \pm \prod_{j=1}^{n} r_{jj} \neq 0,$$

  εφόσον $QQ^T = I_n \;\Rightarrow\; \det Q = \pm 1$ και άρα

  $$A^{-1} = (QR)^{-1} = R^{-1} Q^T.$$

Listing 1: Matlab function file

```
% QR factorization using Gram-Schmidt algorithm
% Input: A is an m-by-n matrix of full rank m>=n
% Output: an m-by-n unitary matrix Q and
%         an n-by-n upper triangular matrix R such that A = Q*R.

function [Q,R] = Gram_Schmidt(A)

[m,n] = size(A);
R = zeros(n);
R(1,1) = norm(A(:,1));
Q(:,1) = A(:,1)/R(1,1);
for k=2:n
    R(1:k-1,k) = Q(:,1:k-1)'*A(:,k);
    Q(:,k) = A(:,k) - Q(:,1:k-1)*R(1:k-1,k);
    R(k,k) = norm(Q(:,k));
    Q(:,k) = Q(:,k)/R(k,k);
end
```

Παράδειγμα

Έστω $\mathbf{u}_1 = (1,1,1)$, $\mathbf{u}_2 = (0,1,1)$, $\mathbf{u}_3 = (0,0,1)$ μια βάση του $\mathbb{R}^3$. Χρησιμοποιώντας τη μέθοδο Gram-Schmidt θα κατασκευάσουμε μια ορθοκανονική βάση.

- Ορθογωνοποίηση βάσης

$$
\begin{aligned}
\mathbf{v}_1 &= (1,1,1) \\
\mathbf{v}_2 &= (0,1,1) - \frac{\langle (0,1,1),(1,1,1)\rangle}{\langle (1,1,1),(1,1,1)\rangle}(1,1,1) = (0,1,1) - \frac{2}{3}(1,1,1) \\
&= \left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right) \\
\mathbf{v}_3 &= (0,0,1) - \frac{\langle (0,0,1),(1,1,1)\rangle}{\langle (1,1,1),(1,1,1)\rangle}(1,1,1) \\
&\qquad - \frac{\langle (0,0,1), \left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right)\rangle}{\langle \left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right), \left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right)\rangle}\left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right) \\
&= (0,0,1) - \frac{1}{3}(1,1,1) - \frac{1}{2}\left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right) = \left(0, -\frac{1}{2}, \frac{1}{2}\right).
\end{aligned}
$$

Παράδειγμα (συνέχεια)

- <u>Κανονικοποίηση βάσης</u>

$$\mathbf{q}_1 = \left(\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}\right), \mathbf{q}_2 = \left(-\frac{\sqrt{6}}{3}, \frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{6}\right), \mathbf{q}_3 = \left(0, -\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right).$$

$$A = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 + \frac{2}{3}\mathbf{v}_1 & \mathbf{v}_3 + \frac{1}{2}\mathbf{v}_2 + \frac{1}{3}\mathbf{v}_1 \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} \begin{bmatrix} 1 & \frac{2}{3} & \frac{1}{3} \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix} = VT$$

Αν $D = \begin{bmatrix} \|\mathbf{v}_1\| & 0 & 0 \\ 0 & \|\mathbf{v}_2\| & 0 \\ 0 & 0 & \|\mathbf{v}_3\| \end{bmatrix} = \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & \frac{\sqrt{6}}{3} & 0 \\ 0 & 0 & \frac{\sqrt{2}}{2} \end{bmatrix}$, τότε $A = \underbrace{VD^{-1}}_{Q}\underbrace{DT}_{R}$, όπου

$$Q = \begin{bmatrix} \frac{\sqrt{3}}{3} & -\frac{\sqrt{6}}{3} & 0 \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{6} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} \end{bmatrix} \text{ και } R = \begin{bmatrix} \sqrt{3} & \frac{2\sqrt{3}}{3} & \frac{\sqrt{3}}{3} \\ 0 & \frac{\sqrt{6}}{3} & \frac{\sqrt{6}}{6} \\ 0 & 0 & \frac{\sqrt{2}}{2} \end{bmatrix}.$$

Σημειώνουμε οτι $Q^*Q = I_3$ και R αντιστρέψιμος.

Να κατασκευαστεί η απλή QR παραγοντοποίηση του πίνακα

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 2 & 1 \\ 2 & -1 & 2 \\ 1 & 1 & 1 \\ 2 & 0 & -2 \end{bmatrix} \in \mathbb{R}^{5\times 3}.$$

1. Ο βαθμός του πίνακα είναι $rank(A) = 3$ (full rank).
2. Έστω $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix}$. Εφαρμόζουμε τη μέθοδο Gram-Schmidt για να κατασκευάσουμε ορθογώνια διανύσματα

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{a}_1 = \begin{bmatrix} 1 & 1 & 2 & 1 & 2 \end{bmatrix}^T \\ \mathbf{v}_2 &= \mathbf{a}_2 - \frac{0}{11}\mathbf{v}_1 = \begin{bmatrix} -1 & 2 & -1 & 1 & 0 \end{bmatrix}^T \\ \mathbf{v}_3 &= \mathbf{a}_3 - \frac{3}{11}\mathbf{v}_1 - \frac{0}{7}\mathbf{v}_2 = \begin{bmatrix} \frac{8}{11} & \frac{8}{11} & \frac{16}{11} & \frac{8}{11} & -\frac{28}{11} \end{bmatrix}^T. \end{aligned}$$

3. Λύνουμε τις σχέσεις του βήματος 2 ως προς τις στήλες του πίνακα A

$$\begin{aligned}\mathbf{a}_1 &= \mathbf{v}_1\\ \mathbf{a}_2 &= 0\mathbf{v}_1 + \mathbf{v}_2\\ \mathbf{a}_3 &= \frac{3}{11}\mathbf{v}_1 + 0\mathbf{v}_2 + \mathbf{v}_3.\end{aligned}$$

Συνεπώς,

$$A = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} \begin{bmatrix} 1 & 0 & \frac{3}{11} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \underbrace{ED^{-1}}_{Q}\underbrace{DT}_{R} = \begin{bmatrix} \frac{1}{\sqrt{11}} & \frac{-1}{\sqrt{7}} & \frac{2}{\sqrt{77}} \\ \frac{1}{\sqrt{11}} & \frac{2}{\sqrt{7}} & \frac{2}{\sqrt{77}} \\ \frac{2}{\sqrt{11}} & \frac{-1}{\sqrt{7}} & \frac{4}{\sqrt{77}} \\ \frac{1}{\sqrt{11}} & \frac{1}{\sqrt{7}} & \frac{2}{\sqrt{77}} \\ \frac{2}{\sqrt{11}} & 0 & \frac{7}{\sqrt{77}} \end{bmatrix} \begin{bmatrix} \sqrt{11} & 0 & \frac{3\sqrt{11}}{11} \\ 0 & \sqrt{7} & 0 \\ 0 & 0 & \frac{12\sqrt{77}}{121} \end{bmatrix},$$

όπου $D = \begin{bmatrix} \|\mathbf{v}_1\| & 0 & 0 \\ 0 & \|\mathbf{v}_2\| & 0 \\ 0 & 0 & \|\mathbf{v}_3\| \end{bmatrix} = \begin{bmatrix} \sqrt{11} & 0 & 0 \\ 0 & \sqrt{7} & 0 \\ 0 & 0 & \frac{4\sqrt{77}}{11} \end{bmatrix}.$

QR Παραγοντοποίηση στο MatLab

Η εντολή

$$[Q, R] = qr(A, 0); \quad \% \text{ compute economy QR factorization}$$

κατασκευάζει την απλή παραγοντοποίηση QR ενός πλήρους βαθμού πίνακα $A \in \mathbb{R}^{m\times n}$ $(m \geq n)$ με πίνακα $Q \in \mathbb{R}^{m\times n}$, $Q^T Q = I_n$ και άνω τριγωνικό πίνακα $R \in \mathbb{R}^{n\times n}$.

Ειδικότερα, στην άσκηση 1 έχουμε:

$$Q = \begin{bmatrix} -0.3015 & 0.3780 & -0.2279 \\ -0.3015 & -0.7559 & -0.2279 \\ -0.6030 & 0.3780 & -0.4558 \\ -0.3015 & -0.3780 & -0.2279 \\ -0.6030 & 0 & 0.7977 \end{bmatrix}, \quad R = \begin{bmatrix} -3.3166 & 0 & -0.9045 \\ 0 & -2.6458 & 0 \\ 0 & 0 & -3.1909 \end{bmatrix}.$$

Άσκηση 2

Να βρεθεί η απλή QR του πίνακα $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$.

Έστω $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ οι στήλες του A, εφόσον $rank(A) = 3$. Τότε με τη μέθοδο Gram-Schmidt βρίσκουμε τα ορθογώνια διανύσματα

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{a}_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}^T \\ \mathbf{v}_2 &= \mathbf{a}_2 - \frac{3}{4}\mathbf{v}_1 = \begin{bmatrix} -\frac{3}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}^T \\ \mathbf{v}_3 &= \mathbf{a}_3 - \frac{1}{2}\mathbf{v}_1 - \frac{2}{3}\mathbf{v}_2 = \begin{bmatrix} 0 & -\frac{2}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}^T. \end{aligned}$$

$$A = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} \begin{bmatrix} 1 & \frac{3}{4} & \frac{1}{2} \\ 0 & 1 & \frac{2}{3} \\ 0 & 0 & 1 \end{bmatrix} = \underbrace{VD^{-1}}_{Q}\underbrace{DT}_{R} = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{1}{2} & \frac{1}{2\sqrt{3}} & -\sqrt{\frac{2}{3}} \\ \frac{1}{2} & \frac{1}{2\sqrt{3}} & \frac{1}{\sqrt{6}} \\ \frac{1}{2} & \frac{1}{2\sqrt{3}} & \frac{1}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} 2 & \frac{3}{2} & 1 \\ 0 & \frac{\sqrt{3}}{2} & \frac{1}{\sqrt{3}} \\ 0 & 0 & \sqrt{\frac{2}{3}} \end{bmatrix},$$

όπου $D = \begin{bmatrix} \|\mathbf{v}_1\| & 0 & 0 \\ 0 & \|\mathbf{v}_2\| & 0 \\ 0 & 0 & \|\mathbf{v}_3\| \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & \sqrt{\frac{2}{3}} \end{bmatrix}$.