Matrix methods – QR factorisation (1)

MATH 2740 – Mathematics of Data Science – Lecture 07

Julien Arino

julien.arino@umanitoba.ca

Department of Mathematics @ University of Manitoba

Fall 202X

The University of Manitoba campuses are located on original lands of Anishinaabeg, Ininew, Anisininew, Dakota and Dene peoples, and on the National Homeland of the Red River Métis. We respect the Treaties that were made on these territories, we acknowledge the harms and mistakes of the past, and we dedicate ourselves to move forward in partnership with Indigenous communities in a spirit of Reconciliation and collaboration.

Outline

Matrix factorisations

Orthogonality and projections

Gram-Schmidt orthogonalisation process

Orthogonal matrices

Matrix factorisations

Orthogonality and projections

Gram-Schmidt orthogonalisation process

Orthogonal matrices

Matrix factorisations

Matrix factorisations are popular because they allow to perform some computations more easily

There are several different types of factorisations. Here, we study just the QR factorisation, which is useful for many least squares problems

Matrix factorisations

Orthogonality and projections

Gram-Schmidt orthogonalisation process

Orthogonal matrices

Definition 57 (Orthogonal set of vectors)

The set of vectors $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\} \in \mathbb{R}^n$ is an **orthogonal set** if

$$\forall i, j = 1, \ldots, k, \quad i \neq j \implies \boldsymbol{v}_i \bullet \boldsymbol{v}_j = 0$$

Theorem 58

$\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\} \in \mathbb{R}^n$ with $\forall i$, $\boldsymbol{v}_i \neq \boldsymbol{0}$, orthogonal set $\implies \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\} \in \mathbb{R}^n$ linearly independent

Definition 59 (Orthogonal basis)

Let S be a basis of the subspace $W \subset \mathbb{R}^n$ composed of an orthogonal set of vectors. We say S is an **orthogonal basis** of W

Proof of Theorem 58

Assume $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ orthogonal set with $\boldsymbol{v}_i \neq \boldsymbol{0}$ for all $i = 1, \ldots, k$. Recall $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ is LI if

$$c_1 \boldsymbol{v}_1 + \cdots + c_k \boldsymbol{v}_k = \boldsymbol{0} \iff c_1 = \cdots = c_k = 0$$

So assume $c_1, \ldots, c_k \in \mathbb{R}$ are s.t. $c_1 \boldsymbol{v}_1 + \cdots + c_k \boldsymbol{v}_k = \boldsymbol{0}$. Recall that $\forall \boldsymbol{x} \in \mathbb{R}^k$, $\boldsymbol{0}_k \bullet \boldsymbol{x} = 0$. So for some $\boldsymbol{v}_i \in \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$

$$\begin{aligned} 0 &= \boldsymbol{0} \bullet \boldsymbol{v}_i \\ &= (c_1 \boldsymbol{v}_1 + \cdots + c_k \boldsymbol{v}_k) \bullet \boldsymbol{v}_i \\ &= c_1 \boldsymbol{v}_1 \bullet \boldsymbol{v}_i + \cdots + c_k \boldsymbol{v}_k \bullet \boldsymbol{v}_i \end{aligned} \tag{1}$$

As $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ orthogonal, $\boldsymbol{v}_j \bullet \boldsymbol{v}_i = 0$ when $i \neq j$, (1) reduces to

$$c_i \boldsymbol{v}_i \bullet \boldsymbol{v}_i = 0 \iff c_i \|\boldsymbol{v}_i\|^2 = 0$$

As $\boldsymbol{v}_i \neq 0$ for all i, $\|\boldsymbol{v}_i\| \neq 0$ and so $c_i = 0$. This is true for all i, hence the result $\square$

Example – Vectors of the standard basis of $\mathbb{R}^3$

For $\mathbb{R}^3$, we denote

$$\boldsymbol{i} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{j} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \text{ and } \boldsymbol{k} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

($\mathbb{R}^k$ for $k > 3$, we denote them $\boldsymbol{e}_i$)
Clearly, $\{\boldsymbol{i}, \boldsymbol{j}\}$, $\{\boldsymbol{i}, \boldsymbol{k}\}$, $\{\boldsymbol{j}, \boldsymbol{k}\}$ and $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ orthogonal sets. The standard basis vectors are also $\neq \mathbf{0}$, so the sets are LI. And

$$\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$$

is an orthogonal basis of $\mathbb{R}^3$ since it spans $\mathbb{R}^3$ and is LI

$$c_1\boldsymbol{i} + c_2\boldsymbol{j} + c_3\boldsymbol{k} = c_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + c_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

Orthonormal version of things

Definition 60 (Orthonormal set)

The set of vectors $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\} \in \mathbb{R}^n$ is an **orthonormal set** if it is an orthogonal set and furthermore

$$\forall i = 1, \ldots, k, \quad \|\boldsymbol{v}_i\| = 1$$

Definition 61 (Orthonormal basis)

A basis of the subspace $W \subset \mathbb{R}^n$ is an **orthonormal basis** if the vectors composing it are an orthonormal set

$\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\} \in \mathbb{R}^n$ is orthonormal if

$$\boldsymbol{v}_i \bullet \boldsymbol{v}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

Projections

Definition 62 (Orthogonal projection onto a subspace)

$W \subset \mathbb{R}^n$ a subspace and $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ an orthogonal basis of W. $\forall \boldsymbol{v} \in \mathbb{R}^n$, the **orthogonal projection** of $\boldsymbol{v}$ **onto** W is

$$\mathrm{proj}_W(\boldsymbol{v}) = \frac{\boldsymbol{u}_1 \bullet \boldsymbol{v}}{\|\boldsymbol{u}_1\|^2}\boldsymbol{u}_1 + \cdots + \frac{\boldsymbol{u}_k \bullet \boldsymbol{v}}{\|\boldsymbol{u}_k\|^2}\boldsymbol{u}_k$$

Definition 63 (Component orthogonal to a subspace)

$W \subset \mathbb{R}^n$ a subspace and $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ an orthogonal basis of W. $\forall \boldsymbol{v} \in \mathbb{R}^n$, the **component** of $\boldsymbol{v}$ **orthogonal to** W is

$$\mathrm{perp}_W(\boldsymbol{v}) = \boldsymbol{v} - \mathrm{proj}_W(\boldsymbol{v})$$

Matrix factorisations

Orthogonality and projections

Gram-Schmidt orthogonalisation process

Orthogonal matrices

What this aims to do is to construct an orthogonal basis for a subspace $W \subset \mathbb{R}^n$

To do this, we use the *Gram-Schmidt orthogonalisation process*, which turn s a basis of W into an orthogonal basis of W

Gram-Schmidt process

Theorem 64

$W \subset \mathbb{R}^n$ a subset and $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}$ a basis of W. Let

$$\begin{aligned}
\boldsymbol{v}_1 &= \boldsymbol{x}_1 \\
\boldsymbol{v}_2 &= \boldsymbol{x}_2 - \frac{\boldsymbol{v}_1 \bullet \boldsymbol{x}_2}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 \\
\boldsymbol{v}_3 &= \boldsymbol{x}_3 - \frac{\boldsymbol{v}_1 \bullet \boldsymbol{x}_3}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 - \frac{\boldsymbol{v}_2 \bullet \boldsymbol{x}_3}{\|\boldsymbol{v}_2\|^2}\boldsymbol{v}_2 \\
&\vdots \\
\boldsymbol{v}_k &= \boldsymbol{x}_k - \frac{\boldsymbol{v}_1 \bullet \boldsymbol{x}_k}{\|\boldsymbol{v}_1\|^2}\boldsymbol{v}_1 - \cdots - \frac{\boldsymbol{v}_{k-1} \bullet \boldsymbol{x}_k}{\|\boldsymbol{v}_{k-1}\|^2}\boldsymbol{v}_{k-1}
\end{aligned}$$

and

$$W_1 = \operatorname{span}(\boldsymbol{x}_1), W_2 = \operatorname{span}(\boldsymbol{x}_1, \boldsymbol{x}_2), \ldots, W_k = \operatorname{span}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k)$$

Then $\forall i = 1, \ldots, k$, $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i\}$ orthogonal basis for W_i

Matrix factorisations

Orthogonality and projections

Gram-Schmidt orthogonalisation process

Orthogonal matrices

Theorem 65

Let $Q \in \mathcal{M}_{mn}$. The columns of Q form an orthonormal set if and only if

$$Q^T Q = \mathbb{I}_n$$

Definition 66 (Orthogonal matrix)

$Q \in \mathcal{M}_n$ is an **orthogonal matrix** if its columns form an orthonormal set

So $Q \in \mathcal{M}_n$ orthogonal if $Q^T Q = \mathbb{I}$, i.e., $Q^T = Q^{-1}$

Theorem 67 (NSC for orthogonality)

$Q \in \mathcal{M}_n$ orthogonal $\iff Q^{-1} = Q^T$

Theorem 68 (Orthogonal matrices “encode" isometries)

Let $Q \in \mathcal{M}_n$. TFAE

1. *Q orthogonal*
2. $\forall \boldsymbol{x} \in \mathbb{R}^n$, $\|Q\boldsymbol{x}\| = \|\boldsymbol{x}\|$
3. $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, $Q\boldsymbol{x} \bullet Q\boldsymbol{y} = \boldsymbol{x} \bullet \boldsymbol{y}$

Theorem 69

Let $Q \in \mathcal{M}_n$ be orthogonal. Then

1. *The rows of Q form an orthonormal set*
2. *Q^{-1} orthogonal*
3. $\det Q = \pm 1$
4. $\forall \lambda \in \sigma(Q)$, $|\lambda| = 1$
5. *If $Q_2 \in \mathcal{M}_n$ also orthogonal, then QQ_2 orthogonal*

Proof of 4 in Theorem 69

All statements in Theorem 69 are easy, but let's focus on 4

Let λ be an eigenvalue of $Q \in \mathcal{M}_n$ orthogonal, i.e., $\exists \mathbb{R}^n \ni \boldsymbol{x} \neq \boldsymbol{0}$ s.t.

$$Q\boldsymbol{x} = \lambda \boldsymbol{x}$$

Take the norm on both sides

$$\|Q\boldsymbol{x}\| = \|\lambda \boldsymbol{x}\|$$

From 2 in Theorem 68, $\|Q\boldsymbol{x}\| = \|\boldsymbol{x}\|$ and from the properties of norms, $\|\lambda \boldsymbol{x}\| = |\lambda| \, \|\boldsymbol{x}\|$, so we have

$$\|Q\boldsymbol{x}\| = \|\lambda \boldsymbol{x}\| \iff \|\boldsymbol{x}\| = |\lambda| \, \|\boldsymbol{x}\| \iff 1 = |\lambda|$$

(we can divide by $\|\boldsymbol{x}\|$ since $\boldsymbol{x} \neq \boldsymbol{0}$ as an eigenvector) $\square$