Orthogonalisation

Sources:

1. `Matrix Methods for Computational Modeling and Data Analytics` author: Mark Embree, Virginia Tech

2. `Linear Algebra : Theory, Intuition, Code` author: Mike X Cohen, publisher: sincXpress

Motivation

Orthogonal Basis

A set of vectors $\{\mathbf{q}_1, \dots,\ \mathbf{q}_n\}$ is a orthonormal if these conditions are met:

1. Vectors are mutually orthogonal : $\mathbf{q}_j^T \cdot \mathbf{q}_k = 0\ :\ j \neq k$

2. $||\mathbf{q}_j|| = 1 :\ j = 1, \dots,\ n$

Using the set of orthonormal vectors $\{\mathbf{q}_1, \dots,\ \mathbf{q}_n\}$ a matrix $\mathbf{Q}$ is defined.

$$\mathbf{Q} = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \\ | & | & \cdots & | \end{bmatrix}$$

The matrix product $\mathbf{Q}^T \cdot \mathbf{Q}$ has elements which are the inner products between vectors $\mathbf{q}_j$ and $\mathbf{q}_k$.

$$\mathbf{Q}^T \cdot \mathbf{Q} = \begin{bmatrix} \mathbf{q}_1^T \cdot \mathbf{q}_1 & \mathbf{q}_1^T \cdot \mathbf{q}_2 & \cdots & \mathbf{q}_1^T \cdot \mathbf{q}_n \\ \mathbf{q}_2^T \cdot \mathbf{q}_1 & \mathbf{q}_2^T \cdot \mathbf{q}_2 & \ddots & \mathbf{q}_2^T \cdot \mathbf{q}_n \\ \vdots & \ddots & \ddots & \mathbf{q}_{n-1}^T \cdot \mathbf{q}_n \\ \mathbf{q}_n^T \cdot \mathbf{q}_1 & \cdots & \mathbf{q}_n^T \cdot \mathbf{q}_{n-1} & \mathbf{q}_n^T \cdot \mathbf{q}_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix} = \mathbf{I}$$

Definition / unitary matrix

A $n \times n$ matrix $\mathbf{Q}$ is *unitary* if $\mathbf{Q}^T \cdot \mathbf{Q} = \mathbf{I} :\ \ \in \mathrm{R}^{n \times n}$

Definition / sub-unitary matrix

A $m \times n\ :\ m > n$ matrix $\mathbf{Q}$ is *sub-unitary* if $\mathbf{Q}^T \cdot \mathbf{Q} = \mathbf{I} :\ \ \in \mathrm{R}^{m \times m}$

Problem#1

For $\mathbf{Q} \in \mathrm{R}^{m \times n}$ and $m < n$: is $\mathbf{Q}^T \cdot \mathbf{Q} = \mathbf{I}$ possible ?

No: it is only possible to have $m < n$ lineary independent column vectors of $\mathbf{Q}$.

Problem#2

Let $\mathbf{Q} \in \mathrm{R}^{m \times n} \;:\; m > n$ be *sub-unitary*. Compute $\mathbf{Q} \cdot \mathbf{Q}^T \in \mathrm{R}^{m \times m}$.

Using the *layer perspective* of matrix multiplication the product of two matrices $\mathbf{A} \in \mathrm{R}^{m \times n}$ and $\mathbf{B} \in \mathrm{R}^{n \times k}$ can expressed as the sum of `n` sub-matrices (layers). Each submatrix is an outer product of the `j-th` column vector of $\mathbf{A}$ and the `j-th` row vector of $\mathbf{B}$. Applying the layer perspective to the case of $\mathbf{Q} \cdot \mathbf{Q}^T$ yields the following equation:

$$\mathbf{\Pi} = \mathbf{Q} \cdot \mathbf{Q}^T = \sum_{j=1}^{n} \underbrace{\mathbf{q}_j \cdot \mathbf{q}_j^T}_{j'th\ layer}$$

Next we show $\mathbf{\Pi} \cdot \mathbf{\Pi} = \mathbf{\Pi}^2 = \mathbf{\Pi}$:

$$\mathbf{\Pi} \cdot \mathbf{\Pi} = \mathbf{\Pi}^2 = \mathbf{Q} \cdot \mathbf{Q}^T \cdot \mathbf{Q} \cdot \mathbf{Q}^T = \mathbf{Q} \cdot \underbrace{\left(\mathbf{Q}^T \cdot \mathbf{Q}\right)}_{\mathbf{I}} \cdot \mathbf{Q}^T = \mathbf{Q} \cdot \mathbf{Q}^T = \mathbf{\Pi}$$

This property can generalised:

$$\mathbf{\Pi}^p = \mathbf{\Pi}$$

Constructing a orthonormal basis from some other basis (Gram-Schmidt process)

Starting point is a set of basis vectors (in general not orthogonal) $\{\mathbf{a}_1,\ \ldots,\ \mathbf{a}_n\}$ for a subspace $V \subset \mathrm{R}^m$. From this basis an orthonormal basis $\{\mathbf{q}_1,\ \ldots,\ \mathbf{q}_n\}$ for the same subspace shall be constructed.

The construction of orthonormal basis vectors is a step-wise process. Each step generates a new vector orthonormal to all vectors generated in <u>previous</u> steps.

step#1

Generating the first vector $\mathbf{q}_1$ is easy. Just take vector $\mathbf{a}_1$ and normalise it:

$$\mathbf{q}_1 = \frac{\mathbf{a}_1}{||\mathbf{a}_1||}$$

step#2

Vector $\mathbf{a}_2$ is used to generate the next vector $\mathbf{q}_2$ which is orthonormal to $\mathbf{q}_1$. First the part of $\mathbf{a}_2$ which is in the direction of $\mathbf{q}_1$ is removed from $\mathbf{a}_2$. This results in a vector $\mathbf{u}_2$ which is already orthogonal to $\mathbf{u}_1$. However it is not yet orthonormal. Thus $\mathbf{u}_2$ is normalised to give $\mathbf{q}_2$.

$$\mathbf{u}_2 = \mathbf{a}_2 - \frac{\mathbf{a}_2^T \cdot \mathbf{q}_1}{||\mathbf{q}_1^2||} \cdot \mathbf{q}_1$$

$$\mathbf{q}_2 = \frac{\mathbf{u}_2}{||\mathbf{u}_2||}$$

Just to make sure it is shown that vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ are orthogonal. We compute

$$\begin{aligned}
\mathbf{q}_1^T \cdot \mathbf{u}_2 = \mathbf{q}_1^T \cdot (\mathbf{a}_2 - \frac{\mathbf{a}_2^T \cdot \mathbf{q}_1}{||\mathbf{q}_1^2||}) \; &= \mathbf{q}_1^T \cdot \mathbf{a}_2 - \mathbf{q}_1^T \cdot \frac{\mathbf{a}_2^T \cdot \mathbf{q}_1}{||\mathbf{q}_1^2||} \cdot \mathbf{q}_1 &(1)\\
&= \mathbf{q}_1^T \cdot \mathbf{a}_2 - \frac{\mathbf{a}_2^T \cdot \mathbf{q}_1}{||\mathbf{q}_1^2||} \cdot ||\mathbf{q}_1||^2 &(2)\\
&= \mathbf{q}_1^T \cdot \mathbf{a}_2 - \mathbf{a}_2^T \cdot \mathbf{q}_1 = \mathbf{q}_1^T \cdot \mathbf{a}_2 - \mathbf{q}_1^T \cdot \mathbf{a}_2 = 0 &(3)
\end{aligned}$$

step#k

Assuming $k < n$:

Having arrived at this step we have already generated orthonormal basis vectors $\{\mathbf{q}_1,\ \mathbf{q}_2, \ldots,\ \mathbf{q}_{k-1}\}$

The procedure is goes like this:

1. Eliminate the part of $\mathbf{a}_k$ which is in the direction of $\mathbf{q}_1$. The residual vector is orthogonal to $\mathbf{q}_1$.

2. Eliminate the part of the previous residual vector in the direction of $\mathbf{q}_2$. The residual vector is orthogonal to $\mathbf{q}_2$ and $\mathbf{q}_1$

3. Eliminate the part of the previous residual vector in the direction of $\mathbf{q}_3$. The residual vector is orthogonal to $\mathbf{q}_3$ and $\mathbf{q}_2$ and $\mathbf{q}_1$.

Repeat ...

4. Finally eliminate the part of the previous residual vector in the direction of $\mathbf{q}_{k-1}$. The residual vector is orthogonal to $\mathbf{q}_{k-1},\ \ldots,\ \mathbf{q}_2,\ \mathbf{q}_1$.

5. Normalise the residual vector and assign it to $\mathbf{q}_k$.

$$\mathbf{u}_k = \mathbf{a}_k - \sum_{l=1}^{k-1} \frac{\mathbf{a}_k^T \cdot \mathbf{q}_l}{||\mathbf{q}_l||^2} \cdot \mathbf{q}_l \tag{4}$$

$$\rightarrow \; normalise \tag{5}$$

$$\mathbf{q}_k = \frac{1}{||\mathbf{u}_k||} \cdot \mathbf{u}_k \tag{6}$$

The procedure is also known as the `Gram-Schmidt` process.

Example / Gram-Schmidt process

taken from: `Matrix Methods for Computational Modeling and Data Analytics`
author: Mark Embree, Virginia Tech

Three linearly independent vectors $\{\mathbf{a}_1,\ \mathbf{a}_2,\ \mathbf{a}_3\}$ form a basis of a subspace. The basis vectors need not be orthogonal / orthonormal (any kind of basis is sufficient).

Using the `Gram-Schmidt` process we generate an orthonormal basis $\{\mathbf{q}_1,\ \mathbf{q}_2,\ \mathbf{q}_3\}$ for the same subspace.

Computing $\mathbf{q}_1$

$$\mathbf{u}_1 = \mathbf{a}_1 \tag{7}$$

$$\mathbf{q}_1 = \frac{1}{||\mathbf{u}_1||} \cdot \mathbf{u}_1 \tag{8}$$

Computing $\mathbf{q}_2$

$$\mathbf{u}_2 = \mathbf{a}_2 - \frac{\mathbf{a}_2^T \cdot \mathbf{q}_1}{||\mathbf{q}_1||^2} \cdot \mathbf{q}_1 \tag{9}$$

$$\mathbf{q}_2 = \frac{1}{||\mathbf{u}_2||} \cdot \mathbf{u}_2 \tag{10}$$

Computing $\mathbf{q}_3$

$$\mathbf{u}_3 = \mathbf{a}_3 - \frac{\mathbf{a}_3^T \cdot \mathbf{q}_1}{||\mathbf{q}_1||^2} \cdot \mathbf{q}_1 - \frac{\mathbf{a}_3^T \cdot \mathbf{q}_2}{||\mathbf{q}_2||^2} \cdot \mathbf{q}_2 \tag{11}$$

$$\mathbf{q}_3 = \frac{1}{||\mathbf{u}_3||} \cdot \mathbf{u}_3 \tag{12}$$

Having computed the orthonormal basis $\{\mathbf{q}_1,\ \mathbf{q}_2,\ \mathbf{q}_3\}$ we now express vectors $\{\mathbf{a}_1,\ \mathbf{a}_2,\ \mathbf{a}_3\}$ by these orthonormal vectors:

$$\mathbf{a}_1 = ||\mathbf{u}_1|| \cdot \mathbf{q}_1 \tag{13}$$

$$\mathbf{a}_2 = \frac{\mathbf{a}_2^T \cdot \mathbf{q}_1}{||\mathbf{q}_1||^2} \cdot \mathbf{q}_1 + ||\mathbf{u}_2|| \cdot \mathbf{q}_2 \tag{14}$$

$$\mathbf{a}_3 = \frac{\mathbf{a}_3^T \cdot \mathbf{q}_1}{||\mathbf{q}_1||^2} \cdot \mathbf{q}_1 + \frac{\mathbf{a}_3^T \cdot \mathbf{q}_2}{||\mathbf{q}_2||^2} \cdot \mathbf{q}_2 + ||\mathbf{u}_3|| \cdot \mathbf{q}_3 \tag{15}$$

Let vectors $\{\mathbf{a}_1,\ \mathbf{a}_2,\ \mathbf{a}_3\}$ be the column vectors of a matrix $\mathbf{A}$ and vectors $\{\mathbf{q}_1,\ \mathbf{q}_2,\ \mathbf{q}_3\}$ be the column vectors of a matrix $\mathbf{Q}$.

$$\mathbf{A} = \begin{bmatrix} | & | & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \\ | & | & | \end{bmatrix}$$

$$\mathbf{Q} = \begin{bmatrix} | & | & | \\ \mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3 \\ | & | & | \end{bmatrix}$$

We may then express matrix $\mathbf{A}$ by multiplying matrix $\mathbf{Q}$ from the right by another matrix $\mathbf{R}$.

$$\mathbf{A} = \begin{bmatrix} | & | & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \\ | & | & | \end{bmatrix} = \underbrace{\begin{bmatrix} | & | & | \\ \mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3 \\ | & | & | \end{bmatrix}}_{\mathbf{Q}} \cdot \underbrace{\begin{bmatrix} ||\mathbf{u}_1|| & \frac{\mathbf{a}_2^T \cdot \mathbf{q}_1}{||\mathbf{q}_1||^2} & \frac{\mathbf{a}_3^T \cdot \mathbf{q}_1}{||\mathbf{q}_1||^2} \\ 0 & ||\mathbf{u}_2|| & \frac{\mathbf{a}_3^T \cdot \mathbf{q}_2}{||\mathbf{q}_2||^2} \\ 0 & 0 & ||\mathbf{u}_3|| \end{bmatrix}}_{\mathbf{R}}$$

Note that matrix $\mathbf{R}$ is a upper-triangular matrix.

Definition / QR factorization

The fact that a matrix $\mathbf{A} \in \mathrm{R}^{m \times n}\ \ m \geq n$ with linearly independent columns can be expressed by the product of a subunitary matrix $\mathbf{Q}$ and a upper triangular matrix $\mathbf{R}$ is known as `QR` - decomposition / factorization.

In [ ]: