Chapter 4
Orthogonality

Yuri Lavinas
Master Student - University of Tsukuba

System and Information Engineering

05/25/2018

Why consider orthogonality?

- Gaussian elimination - standard algorithm used in numerous applications.
  - May not be sufficient: Most important data X Less important data.
- Less important - Linearly dependent.
- Most important - close to orthogonal, very linearly. independent.

Example 4.1

- Orthogonal columns determine the plane much better.

Let A and B be matrices in the $\mathbb{R}^3$:

$$A = \begin{pmatrix} 1 & 1.05 \\ 1 & 1 \\ 1 & 0.95 \end{pmatrix},$$

$$B = \begin{pmatrix} 1 & 1/\sqrt(2) \\ 1 & 0 \\ 1 & 1/\sqrt(2) \end{pmatrix}.$$

Figure 4.1

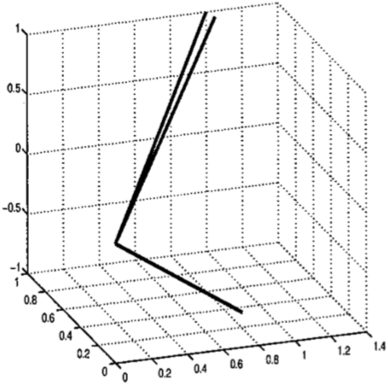

Figure 4.1. *Three vectors spanning a plane in* $\mathbb{R}^3$.

Preposition 4.2

- Two vectors v and u are orthogonal if $x^T y = 0$, $\cos\theta(x, y) = 0$.

Let $q_i, j = 1, 2, ..., n$, *be orthogonal,* $q_i^T q_j = 0$, $i \neq j$. *Then they are linearly independent.*

Proof. Assume the vectors are linearly dependent. From preposition 2.2 there exists a q_k such that

$$q_k = \sum_{j \neq k} \alpha_j q.$$

Proof 4.2

Given

$$q_k = \sum_{j \neq k} \alpha_j q.$$

Multiplying this equation by q_k:

$$q_k^T q_k = \sum_{j \neq k} \alpha_j q_k^T q_j,$$

Since the vectors are orthogonal:

$$q_k^T q_k = \sum_{j \neq k} \alpha_j q_k^T q_j = 0,$$

which is a contradiction (we assumed that they are linealy dependent).

Normalization

Let the set of orthogonal vectors q_j j = 1,2, . . . , m in $\mathbb{R}^M$, be normalized:

$$||q_i||_2 = 1.$$

They are called *orthonormal* and constitute *orthonormal basis* in $\mathbb{R}^M$.

A square matrix

$$Q = (q_1 \;\; q_2 \;\; \cdots \;\; q_m) \in \mathbb{R}^{M\times M}$$

A matrix Q, whose columns are orthonormal, is called orthogonal matrix.

Orthogonal matrices satisfies important properties.

Preposition 4.3 - **Proof.**

- *An orthogonal matrix Q satisfies $Q^T Q = I$*

$$Q^T Q = (q_1 \ \ q_2 \ \ \cdots \ \ q_m)^T (q_1 \ \ q_2 \ \ \cdots \ \ q_m) = \begin{pmatrix} q_1^T \\ q_2^T \\ \vdots \\ q_m^T \end{pmatrix} (q_1 \ \ q_2 \ \ \cdots \ \ q_m)$$

$$\begin{pmatrix} q_1^T q_1 & q_1^T q_2 & \cdots & q_1^T q_m \\ q_2^T q_1 & q_2^T q_2 & \cdots & q_2^T q_m \\ \vdots & \vdots & & \vdots \\ q_m^T q_1 & q_m^T q_2 & \cdots & q_m^T q_m \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix},$$

due to orthogonality.

Proposition 4.4 and 4.5

- *An orthogonal matrix* $Q \in \mathbb{R}^{m\times m}$ *has rank m, and, since* $Q^TQ = I$*, its inverse is equal to* $Q^{-1} = Q^T$*.*
- *The rows of an orthogonal matrix are orthogonal, i.e.,* $QQ^T = I$

Proof.

Proof. Let x be an arbitrary vector. We shall show that $QQ^Tx = x$. Given x there is a uniquely determined vector y, such that $Qy = c$, since Q^{-1} exists. Then

$$QQ^Tx = QQ^TQy = Qy = x.$$

Since x is arbitrary, it follows that $QQ^T = I$.

Proposition 4.6

- *The product of two orthogonal matrices is orthogonal.*

Proof. Let Q and P be orthogonal, and put $X = PQ$. Then

$$X^T X = (PQ)^T PQ = Q^T P^T PQ = Q^T Q = I$$

- Any orthonormal basis of a subspace of $\mathbb{R}^m$ can be enlarged to an orthonormal basis of the whole space.

Proposition 4.7

Given a matrix $Q_1 \in \mathbb{R}^{m\times k}$, with orthonormal columns, there exists a matrix $Q_2 \in \mathbb{R}^{m-k}$ such that $Q = (Q_1 Q_2)$ is an orthogonal matrix.

This is a standard result in linear algebra.

Proposition 4.8

The Euclidean length of a vector is invariant under an orthogonal transformation Q.

Proof. $||Q_x||_2^2 = (QX)^T Qx = x^T Q^T Qx = x^T x = ||x||_2^2$

Also, the corresponding matrix norm and the Frobenius norm are invariant under orthogonal transformations.

Proposition 4.9

Let $U \in \mathbb{R}^{m\times m}$ *and* $V \in \mathbb{R}^{n\times n}$ *be orthogonal. Then for any* $A \in\in \mathbb{R}^{m\times n}$,

$$||UAV||_2 = ||A||_2$$

$$||UAV||_F = ||A||_F$$

Proof. The first equality is proved using Proposition 4.8. For the second, the proof needs the following alternative expression, $||A||_F^2 = tr(A^TA)$, for the Frobenius normal and the identity tr(BC) $=$ tr(CB).

Elementary Orthogonal Matrices

- Use elementary orthogonal matrices to reduce matrices to compact form.
- We will transform a matrix $A \in \mathbb{R}^{m \times n}, m > n$, to triangular form.

Plane rotations

A 2x2 plane rotation matrix

$$G = \begin{pmatrix} c & s \\ -s & c \end{pmatrix}, c^2 + s^2 = 1$$

is orthogonal.

- Multiplying G by a vector x rotates the vector in clockwise direction by a angle θ, where c $= \cos(\theta)$.

Usage

- A plane rotation can be used to zero the second element of a vector x by choosing $c = x1/\sqrt{x_1^2 + x_2^2}$ and $s = x2/\sqrt{x_1^2 + x_2^2}$.

$$\frac{1}{\sqrt{x_1^2 + x_2^2}} \begin{pmatrix} x_1 & x_2 \\ -x_2 & x_1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \sqrt{x_1^2 + x_2^2} \\ 0 \end{pmatrix}$$

It is possible to manipulate vectors and matrices of arbitrary dimension by embedding two-dimensional rotation in a larger matrix.

Example 4.10

We can choose c and s, in

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & c & 0 & s \\ 1 & 0 & 1 & 0 \\ 1 & -s & 0 & c \end{pmatrix}$$

so that the 4th element of the vector $x \in \mathbb{R}^4$ by rotating the plane $(2, 4)$.

Code 4.10

The MATLAB script follows:

```
x = [1;2;3;4];
sq = sqrt(x(2)^2 + x(4)^2);
c = x(2)/sq; s = x(4)/sq;
G = [1 0 0 0; 0 c 0 s; 0 0 1 0; 0 -s 0 c];
y = G * x
```

given the result:

```
y = 1.0000
4.4721
3.0000
0
```

Transform an arbitrary vector

Using a sequence of of planes rotations, is possible to transform any vector to a multiple of a unit vector.
Given a $x \in \mathbb{R}^4$, we transform in to ke_1. First, by a rotation G_3 in the plane (3,4) the last element became zero:

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & s \\ 0 & 0 & c_1 & s_1 \\ 0 & 0 & -s_1 & c_1 \end{pmatrix} \begin{pmatrix} \times \\ \times \\ \times \\ \times \end{pmatrix} = \begin{pmatrix} \times \\ \times \\ * \\ 0 \end{pmatrix}.$$

Next rotation

A rotation G_2 in the plane (2,3) the element in position 3 became zero:

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & c_2 & s_2 & 0 \\ 0 & -s_2 & c_2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \times \\ \times \\ \times \\ 0 \end{pmatrix} = \begin{pmatrix} \times \\ * \\ 0 \\ 0 \end{pmatrix}.$$

Final rotation

And finally, for the second element

$$G = \begin{pmatrix} c_3 & s_3 & 0 & 0 \\ -s_3 & c_3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \times \\ \times \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} K \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Summary

- According to proposition 4.8 the Euclidean length is preserved, and therefore we know that k $= ||x||_2$.
- Since the product of orthogonal matrices is orthogonal (Proposition 5.6) the matrix $P = G_1G_2G_3$ is orthogonal, and the overall results is $Px = ke_1$.
- Rotations are very flexible and can be used efficiently for problems with a sparsity structure, e.g., band matrices.
- For dense matrices they require more flops than Householder transformations (section 4.3).

Example 4.12

- In the last MATLAB example, a 2 x 2 matrix was explicit embedded in a matrix of larger dimension. This is a waste of operations, since the execution of the code does not consider that only two rows are changed.
- The whole matrix multiplication is performed - $2n^3$ flops, for a matrix of dimension n.
- The following MATLAB code illustrates an alternative approach to save operations (and storage).

Code 4.12

The MATLAB script follows:

```
function [c, s] = rot(x, y);
        sq = sqrt(x^2 + y^2);
        c = x/sq; s= y/sq;

function X = approt(c, s, i, ,j, X);
        X([i, j],:) = [c s; -s c]*X([i,j],:);
```

Code 4.12

Continuing:

```
x[1;2;3;4];
for i=3:-1:1
    [c,s] = rot(x(i),x(i+1));
end;

x = 5.4772
       0
       0
       0
```

- After the reduction the first component of x is equal to $||x||_2$.

Table of Contents

Orthogonality
- Introduction
- Section 4.1
- Section 4.2