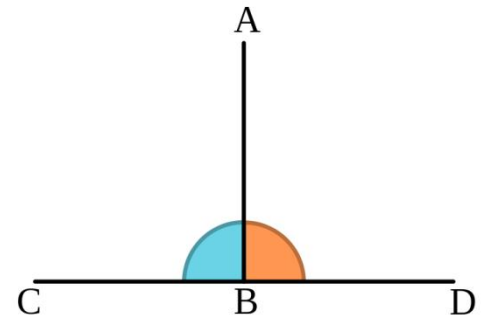# 3 Orthogonality
（正交性）

3.4

ORTHOGONAL BASES AND GRAM-SCHMIDT
（正交基；Gram-Schmidt正交化）

Orthogonal Matrices

Rectangular Matrices with Orthonormal Columns

Gram-Schmidt Orthogonalization

QR factorization

For $\mathbf{R}^n$, the standard basis $\{e_1, e_2, ..., e_n\}$ satisfies

$\qquad e_i^T e_j = 1$ or 0 depending on $i = j$ or $i \neq j$, respectively.

They are called orthonormal.

In general, the vectors $q_1, q_2, ..., q_n$ are called **orthonormal** (标准正交, 单位正交) if

$$q_i^T q_j = \begin{cases} 0 & \text{whenever } i \neq j, \\ 1 & \text{whenever } i = j. \end{cases}$$

giving the orthogonality; (两两正交)

giving the normalization. (单位长度)

We will see:

1. What happens with a matrix with orthonormal columns – square (*orthogonal matrix*, 正交矩阵) or rectangular;

2. A subspace always has an orthonormal basis, and it can be constructed in a simple way out of any basis whatsoever.

· the method *converting a skewed set of axes into a perpendicular set* – *Gram-Schmidt orthogonalization* (Gram-Schmidt正交化).

I. Orthogonal Matrices (正交矩阵)

Definition 1 If Q (square or rectangular) has orthonormal columns, then $Q^{\mathrm{T}}Q = I$:

$$\begin{bmatrix} — & q_1^{\mathrm{T}} & — \\ — & q_2^{\mathrm{T}} & — \\ & \vdots & \\ — & q_n^{\mathrm{T}} & — \end{bmatrix} \begin{bmatrix} | & | & & | \\ q_1 & q_2 & \cdots & q_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdot & 0 \\ 0 & 1 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & 1 \end{bmatrix} = I.$$

An orthogonal matrix is a *square* *matrix with orthonormal columns.*

(正交矩阵是方阵, 它的列是标准正交的向量组)

Then Q^{T} is Q^{-1}.

(For orthogonal matrices, *the transpose is the inverse.*)

Notes:

An orthogonal matrix is square. (正交矩阵特指满足上述条件的方阵)

$Q^{\mathrm{T}}Q = I$ even if Q is rectangular. But then Q^{T} is only a left-inverse.

Example 1

$$Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \qquad Q^T = Q^{-1} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}.$$

Q is an orthogonal matrix.   *Rotation*

The matrix Q^T is just as much an orthogonal matrix as Q.

Example 2

Any permutation matrix is an orthogonal matrix.

$$P_1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \qquad P_1^{-1} = P_1^T = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

$$P_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \text{ is also an orthogonal matrix and takes } \begin{bmatrix} x \\ y \\ z \end{bmatrix} \text{ to } \begin{bmatrix} z \\ y \\ x \end{bmatrix}.$$

$$P_3 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \text{ reflects } \begin{bmatrix} x \\ y \end{bmatrix} \text{ into } \begin{bmatrix} y \\ x \end{bmatrix}. \qquad \textit{Reflection}$$

Example 3   Decide whether the following matrices are orthogonal matrices:

(1) $\begin{bmatrix} 1 & -1/2 & 1/3 \\ -1/2 & 1 & 1/2 \\ 1/3 & 1/2 & -1 \end{bmatrix}$;   ✗   **(2)** $\begin{bmatrix} 1/9 & -8/9 & -4/9 \\ -8/9 & 1/9 & -4/9 \\ -4/9 & -4/9 & 7/9 \end{bmatrix}$;   ✓

(3) $\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 & 0 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 & 0 \\ 0 & 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$.   ✓

Theorem 1 (*nice properties of orthogonal matrices*)

Multiplication by any $n \times n$ orthogonal matrix $\boldsymbol{Q}$ preserves lengths:

$$\|\boldsymbol{Qx}\| = \|\boldsymbol{x}\| \text{ for every vector } \boldsymbol{x} \in \mathbf{R}^n.$$

It also preserves inner products and angles.

(向量的内积、长度及向量间的夹角都保持不变)

Proof $\|\boldsymbol{Qx}\| = \|\boldsymbol{x}\|$ because $(\boldsymbol{Qx})^{\mathrm{T}}(\boldsymbol{Qx}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Qx} = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}$.

And since $(\boldsymbol{Qx})^{\mathrm{T}}(\boldsymbol{Qy}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Qy} = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}$, $\boldsymbol{Q}$ also preserves inner products and thus the angles.

Remark: This property is shared by *rotations* and *reflections*, and in fact by every orthogonal matrix. It is not shared by projections, which are not orthogonal (some projections are not even invertible).

Theorem 2 Let A, B be $n \times n$ orthogonal matrices, then

(1) $A^{-1} = A^{\mathrm{T}}$;

(2) $AA^{\mathrm{T}} = A^{\mathrm{T}}A = I$;

(3) A^{-1} (i.e., A^{T}) is also an orthogonal matrix;

(4) AB is also an orthogonal matrix.

Proof

(1) $A^{\mathrm{T}}A = I$, then $A^{-1} = A^{\mathrm{T}}$. (A is full rank)

(2) $A^{-1} = A^{\mathrm{T}} \Rightarrow AA^{\mathrm{T}} = I$.

(3) $(A^{\mathrm{T}})^{\mathrm{T}}A^{\mathrm{T}} = AA^{\mathrm{T}} = I$, so $A^{\mathrm{T}} = A^{-1}$ is also an orthogonal matrix.

That is: ***The rows of a square matrix are orthonormal whenever the columns are.*** (正交矩阵的行向量组也是 $\mathbf{R}^n$ 的标准正交基.)

(4) $(AB)^{\mathrm{T}}(AB) = (B^{\mathrm{T}}A^{\mathrm{T}})(AB) = B^{\mathrm{T}}IB = B^{\mathrm{T}}B = I$,

so AB is also an orthogonal matrix.

If we have a basis, then any vector is a combination of the basis vectors. This is exceptionally simple for an _orthonormal basis_. The problem is _to find the coefficients of the basis vectors_:

Problem: _Write $\boldsymbol{b}$ as a combination_: $\boldsymbol{b} = x_1\boldsymbol{q}_1 + x_2\boldsymbol{q}_2 + \ldots + x_n\boldsymbol{q}_n$.
Try to find the coordinates $x_1, x_2, \ldots, x_n$.

Trick: Multiply both sides of the equation by $\boldsymbol{q}_i^{\mathrm{T}}$.

The only term _survives_ is $x_i\boldsymbol{q}_i^{\mathrm{T}}\boldsymbol{q}_i$. The other terms _die_ of orthogonality. That is,

$$\boldsymbol{q}_i^{\mathrm{T}}\boldsymbol{b} = x_i\boldsymbol{q}_i^{\mathrm{T}}\boldsymbol{q}_i = x_i.$$

Therefore,

$$\boldsymbol{b} = (\boldsymbol{q}_1^{T}\boldsymbol{b})\boldsymbol{q}_1 + (\boldsymbol{q}_2^{T}\boldsymbol{b})\boldsymbol{q}_2 + \ldots + (\boldsymbol{q}_n^{T}\boldsymbol{b})\boldsymbol{q}_n.$$

Geometrically, every vector $\boldsymbol{b}$ is the sum of its one-dimensional projections onto the lines through the $\boldsymbol{q}$'s.
($\boldsymbol{b}$ 在标准正交基 $\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_n$ 下的坐标的第 i 个分量是 $\boldsymbol{q}_i^{\mathrm{T}}\boldsymbol{b}$ ，即 $\boldsymbol{b}$ 在 $\boldsymbol{q}_i$ 上的投影.)

II. Rectangular Matrices with Orthonormal Columns (列向量标准正交的长方形矩阵)

Problem:

Solve $Qx = b$, where Q is an m by n matrix ($m > n$), with n orthonormal vectors q_i as the columns of Q.

Then we cannot expect to solve $Qx = b$ exactly.

We solve it by least squares.

Tip: Orthonormal columns should make the problem simple.

The key is to notice that *we still have $Q^TQ = I$*. (So Q^T is still the *left-inverse* of Q.)

The normal equations: $Q^TQ\hat{x} = Q^Tb$

Therefore, $\hat{x} = Q^Tb.$

Note: When Q is square, then $\hat{x}$ is the exact solution;

If Q is rectangular, then $\hat{x}$ is the least squares solution.

If Q has orthonormal columns, the least-squares problem becomes easy:

$Qx = b$	rectangular system ($m>n$) with no solution for most b.
$Q^{\mathrm{T}}Q\widehat{x} = Q^{\mathrm{T}}b$	normal equation for the best $\widehat{x}$ — in which $Q^{\mathrm{T}}Q = I$.
$\widehat{x} = Q^{\mathrm{T}}b$	$\widehat{x}_i$ is $q_i^{\mathrm{T}}b$.
$p = Q\widehat{x}$	the projection of b is $(q_1^{\mathrm{T}}b)q_1 + (q_2^{\mathrm{T}}b)q_2 + \ldots + (q_n^{\mathrm{T}}b)q_n$.
$p = QQ^{\mathrm{T}}b$	the projection matrix is $P = QQ^{\mathrm{T}}$.

$(P = Q(Q^{\mathrm{T}}Q)^{-1}Q^{\mathrm{T}} = QQ^{\mathrm{T}}$,
is an $m \times m$ projection matrix;
while $Q^{\mathrm{T}}Q$ is an $n \times n$ identity matrix)

Example 4 The following case is simple but typical.

Suppose we project a point $b = (x,y,z)^{\mathrm{T}}$ onto the x-y plane.

Its projection is $p = (x,y,0)^{\mathrm{T}}$,

and this is the sum of the separate projections onto the x- and y-axes:

$$q_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \text{ and } (q_1^{\mathrm{T}}b)q_1 = \begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix}; \quad q_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \text{ and } (q_2^{\mathrm{T}}b)q_2 = \begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix}.$$

The overall projection matrix is

$$P = q_1 q_1^{\mathrm{T}} + q_2 q_2^{\mathrm{T}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \text{and} \quad P\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}.$$

Remark: *Projection onto a plane = sum of projections onto orthonormal q_1 and q_2.*

III. Gram-Schmidt Orthogonalization (Gram-Schmidt正交化)

Convert independent vectors into orthonormal vectors

In $\mathbf{R}^n$, we try to make the independent vectors $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ orthonormal. (由$\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$出发, 构造出一组标准正交向量 $\boldsymbol{q}_1$, $\boldsymbol{q}_2$, $\boldsymbol{q}_3$.)

Let $\boldsymbol{q}_1 = \boldsymbol{a}/\|\boldsymbol{a}\|$ to make it a unit vector.

Find the projection of $\boldsymbol{b}$ onto $\boldsymbol{q}_1$ (which is the direction of $\boldsymbol{a}$):

$$\boldsymbol{\gamma}_{12} = \left(\boldsymbol{q}_1^{\mathrm{T}}\boldsymbol{b}\right)\boldsymbol{q}_1$$

That component has to be subtracted:
Take $\boldsymbol{B} = \boldsymbol{b} - \boldsymbol{\gamma}_{12} = \boldsymbol{b} - \left(\boldsymbol{q}_1^{\mathrm{T}}\boldsymbol{b}\right)\boldsymbol{q}_1$,
and $\boldsymbol{q}_2 = \boldsymbol{B}/\|\boldsymbol{B}\|$.
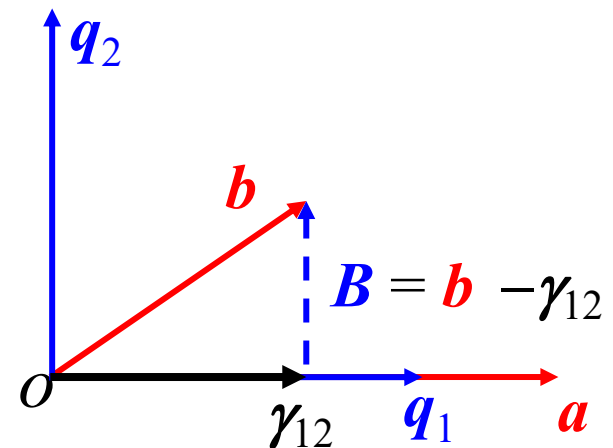
The vector c will not be in the plane of q_1 and q_2, which is the plane of a and b.

However, it may have a component in that plane, and that has to be subtracted:

$$C = c - \gamma_3 = c - (q_1^{\mathrm{T}}c)q_1 - (q_2^{\mathrm{T}}c)q_2,$$

and $q_3 = C/\|C\|$.

When there is a fourth vector in $\mathbf{R}^n$ $(n \geq 4)$, we subtract away its components in the directions of q_1, q_2, q_3.
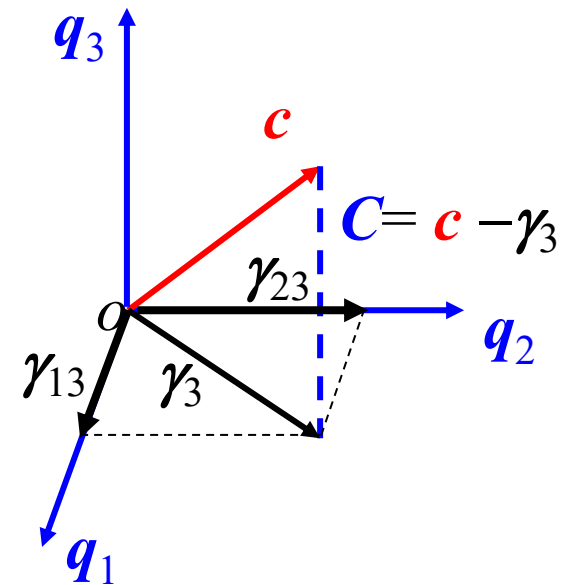
The idea of the Gram-Schmidt process:

to subtract from every new vector its components in the directions that are already settled.

(从每一个新的向量中扣除其在已经确定了的方向上的投影分量)

The Gram-Schmidt process:

starts with *independent* vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, ..., $\boldsymbol{a}_n$ and ends with *orthonormal* vectors $\boldsymbol{q}_1$, $\boldsymbol{q}_2$, ..., $\boldsymbol{q}_n$.

At step j it subtracts from $\boldsymbol{a}_j$ its components in the directions $\boldsymbol{q}_1$, $\boldsymbol{q}_2$, ..., $\boldsymbol{q}_{j-1}$ that are already settled:

$$\boldsymbol{A}_j = \boldsymbol{a}_j - (\boldsymbol{q}_1^{\mathrm{T}}\boldsymbol{a}_j)\boldsymbol{q}_1 - ... - (\boldsymbol{q}_{j-1}^{\mathrm{T}}\boldsymbol{a}_j)\boldsymbol{q}_{j-1}.$$

Then $\boldsymbol{q}_j$ is the unit vector $\boldsymbol{A}_j/\|\boldsymbol{A}_j\|$.

Example 5 Suppose the independent vectors are $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$:

$$\boldsymbol{a} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{c} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}.$$

Then

$$\boldsymbol{q}_1 = \frac{\boldsymbol{a}}{\sqrt{2}};$$

$$\boldsymbol{B} = \boldsymbol{b} - (\boldsymbol{q}_1^{\mathrm{T}} \boldsymbol{b})\boldsymbol{q}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{\sqrt{2}} \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad \boldsymbol{q}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix};$$

$$\boldsymbol{C} = \boldsymbol{c} - (\boldsymbol{q}_1^{\mathrm{T}} \boldsymbol{c})\boldsymbol{q}_1 - (\boldsymbol{q}_2^{\mathrm{T}} \boldsymbol{c})\boldsymbol{q}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} - \sqrt{2} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} - \sqrt{2} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{q}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

Remark:

$\boldsymbol{Q} = [\boldsymbol{q}_1 \quad \boldsymbol{q}_2 \quad \boldsymbol{q}_3] = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$ is an orthogonal matrix

Example 5 Suppose the independent vectors are $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$:

$$\boldsymbol{a} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{c} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}.$$

Remark:

2. It is easier to compute the orthogonal vectors without forcing their lengths to equal one. They can be normalized to unit vector at the end by dividing by their lengths. (先正交化, 再单位化)

$\boldsymbol{A} = \boldsymbol{a}$;

$$\boldsymbol{B} = \boldsymbol{b} - \left(\frac{\boldsymbol{A}^{\mathrm{T}} \boldsymbol{b}}{\boldsymbol{A}^{\mathrm{T}} \boldsymbol{A}} \right) \boldsymbol{A} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix};$$

$$\boldsymbol{C} = \boldsymbol{c} - \left(\frac{\boldsymbol{A}^{\mathrm{T}} \boldsymbol{c}}{\boldsymbol{A}^{\mathrm{T}} \boldsymbol{A}} \right) \boldsymbol{A} - \left(\frac{\boldsymbol{B}^{\mathrm{T}} \boldsymbol{c}}{\boldsymbol{B}^{\mathrm{T}} \boldsymbol{B}} \right) \boldsymbol{B} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - 2 \begin{bmatrix} \frac{1}{2} \\ 0 \\ -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

Then we have $\boldsymbol{q}_1 = \boldsymbol{A}/\|\boldsymbol{A}\|, \quad \boldsymbol{q}_2 = \boldsymbol{B}/\|\boldsymbol{B}\|, \quad \boldsymbol{q}_3 = \boldsymbol{C}/\|\boldsymbol{C}\|.$

IV. QR Factorization (QR分解)

The Gram-Schmidt process:

starts with *independent* vectors a_1, a_2, ..., a_n and ends with *orthonormal* vectors q_1, q_2, ..., q_n.

Let $A = [a_1, a_2, ..., a_n]$,

$$Q = [q_1, q_2, ..., q_n].$$

The matrices A and Q are m by n when the n vectors are in m-dimensional space,

and there has to be a third matrix that connects them: R.
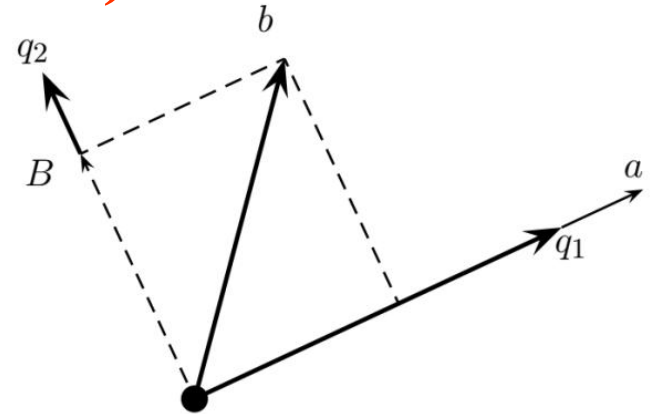
For example,

$A = [a, b, c]$, $Q =$
$[q_1, q_2, q_3]$.

The idea is to write the a, b, c as combinations of the q's:

$a = (q_1^T a) q_1$,

$b = (q_1^T b) q_1 + (q_2^T b) q_2$,

$c = (q_1^T c) q_1 + (q_2^T c) q_2 + (q_3^T c) q_3$.

$$a = \left(q_1^{\mathrm{T}}a\right)q_1,$$
$$b = \left(q_1^{\mathrm{T}}b\right)q_1 + \left(q_2^{\mathrm{T}}b\right)q_2,$$
$$c = \left(q_1^{\mathrm{T}}c\right)q_1 + \left(q_2^{\mathrm{T}}c\right)q_2 + \left(q_3^{\mathrm{T}}c\right)q_3.$$
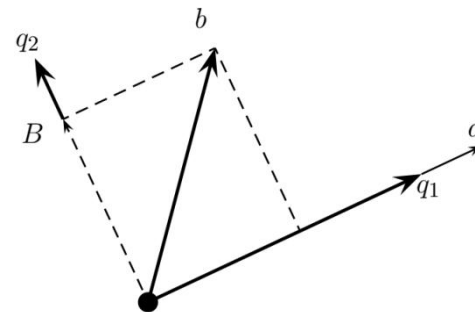
If we express that in matrix form we have ***the new factorization***

A = QR:

$$A = \begin{bmatrix} a & b & c \end{bmatrix} = \begin{bmatrix} q_1 & q_2 & q_3 \end{bmatrix} \begin{bmatrix} q_1^{\mathrm{T}}a & q_1^{\mathrm{T}}b & q_1^{\mathrm{T}}c \\ & q_2^{\mathrm{T}}b & q_2^{\mathrm{T}}c \\ & & q_3^{\mathrm{T}}c \end{bmatrix} = QR.$$

- ***R*** is ~~*upper triangular*~~ because of the way Gram-Schmidt was done. (The first vectors ***a*** and ***q***$_1$ fell on the same line. Then ***q***$_1$, ***q***$_2$ were in the same plane as ***a***, ***b***. The third vectors ***c*** and ***q***$_3$ were not involved until step 3.)

- *You see the lengths of **a**, **B**, **C** on the diagonal of **R**.*

- ***Q*** has orthonormal columns.

Example 5 (Continued) Suppose the independent vectors are a, b, c:

$$a = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad c = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}.$$

And we have $Q = \begin{bmatrix} q_1 & q_2 & q_3 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{bmatrix}.$

The whole factorization is

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{bmatrix} \begin{bmatrix} \sqrt{2} & 1/\sqrt{2} & \sqrt{2} \\ & 1/\sqrt{2} & \sqrt{2} \\ & & 1 \end{bmatrix} = QR.$$

$$\begin{bmatrix} a & b & c \end{bmatrix} = \begin{bmatrix} q_1 & q_2 & q_3 \end{bmatrix} \begin{bmatrix} q_1^{\mathrm{T}}a & q_1^{\mathrm{T}}b & q_1^{\mathrm{T}}c \\ & q_2^{\mathrm{T}}b & q_2^{\mathrm{T}}c \\ & & q_3^{\mathrm{T}}c \end{bmatrix}$$

Note: Another way to find R? $R = Q^{\mathrm{T}}A$.

Theorem 3 Every *m* by *n* matrix *with independent columns* can be factored into $A = QR$. The columns of Q are orthonormal, and R is upper triangular and invertible.

When $m = n$ and all matrices are square, Q becomes an orthogonal matrix. (*Any invertible matrix can be factorized as a product of an orthogonal matrix and an upper triangular matrix.*)

For least-squares problem: $Ax = b$

$A^{\mathrm{T}}A$ becomes easier: $A^{\mathrm{T}}A = R^{\mathrm{T}}Q^{\mathrm{T}}QR = R^{\mathrm{T}}R$.

The normal equation $A^{\mathrm{T}}A\widehat{x} = A^{\mathrm{T}}b$ simplifies to a triangular system: $R^{\mathrm{T}}R\widehat{x} = R^{\mathrm{T}}Q^{\mathrm{T}}b$

$$\Rightarrow \ R\widehat{x} = Q^{\mathrm{T}}b$$

This is just back-substitution because R is triangular.

Key words:

Orthogonal matrices

Rectangular Matrices with Orthonormal Columns

Gram-Schmidt orthogonalization

QR factorization

Homework

See Blackboard