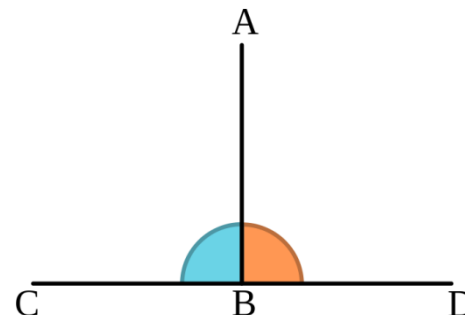


3

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 $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ satisfies

$\mathbf{e}_i^T \mathbf{e}_j = 1$ or 0 depending on $i = j$ or $i \neq j$, respectively.

They are called orthonormal.

In general, the vectors $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n$ are called **orthonormal** (标准正交, 单位正交) if

$$\mathbf{q}_i^T \mathbf{q}_j = \begin{cases} 0 & \text{whenever } i \neq j, \text{ giving the orthogonality; (两两正交)} \\ 1 & \text{whenever } i = j. \text{ giving the normalization. (单位长度)} \end{cases}$$

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.

What method can *convert 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 $Q^T Q = I$:

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

*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**)

$Q^T Q = I$ even if Q is rectangular. But then Q^T is only a **left-inverse**.

Example 1

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

Rotation

\mathbf{Q} is an orthogonal matrix.

The matrix \mathbf{Q}^T is just as much an orthogonal matrix as \mathbf{Q} .

Example 2


Any permutation matrix is an orthogonal matrix.


$$\mathbf{P}_1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \quad \mathbf{P}_1^{-1} = \mathbf{P}_1^T = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$


$$\mathbf{P}_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \text{ is also an orthogonal matrix and takes } \begin{bmatrix} x \\ y \\ z \end{bmatrix} \text{ to } \begin{bmatrix} z \\ y \\ x \end{bmatrix}.$$

$$\mathbf{P}_3 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \text{ reflects } \begin{bmatrix} x \\ y \end{bmatrix} \text{ into } \begin{bmatrix} y \\ x \end{bmatrix}. \quad \text{Reflection}$$

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 Q preserves lengths:

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

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 $\|Q\mathbf{x}\| = \|\mathbf{x}\|$ because $(Q\mathbf{x})^T(Q\mathbf{x}) = \mathbf{x}^T Q^T Q \mathbf{x} = \mathbf{x}^T \mathbf{x}$.

And since $(Q\mathbf{x})^T(Q\mathbf{y}) = \mathbf{x}^T Q^T Q \mathbf{y} = \mathbf{x}^T \mathbf{y}$, 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) $A^{-1} = A^T$;
- (2) $AA^T = A^T A = I$;
- (3) A^{-1} (i.e., A^T) is also an orthogonal matrix;
- (4) AB is also an orthogonal matrix.

Proof

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

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

(3) $(A^T)^T A^T = AA^T = I$, so $A^T = A^{-1}$ 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) $(AB)^T (AB) = (B^T A^T)(AB) = B^T I B = B^T B = I$,

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 \mathbf{b} as a combination: $\mathbf{b} = x_1\mathbf{q}_1 + x_2\mathbf{q}_2 + \dots + x_n\mathbf{q}_n$.

Try to find the coordinates x_1, x_2, \dots, x_n .

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

The only term *survives* is $x_i\mathbf{q}_i^T\mathbf{q}_i$. The other terms *die* of orthogonality. That is,

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

Therefore,

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

Geometrically, every vector \mathbf{b} is the sum of its one-dimensional projections onto the lines through the \mathbf{q} 's.

(\mathbf{b} 在标准正交基 $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n$ 下的坐标的第 i 个分量是 $\mathbf{q}_i^T\mathbf{b}$, 即 \mathbf{b} 在 \mathbf{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 $\mathbf{Q}\mathbf{x} = \mathbf{b}$, where \mathbf{Q} is an m by n matrix ($m > n$), with n orthonormal vectors \mathbf{q}_i as the columns of \mathbf{Q} .

Then we cannot expect to solve $\mathbf{Q}\mathbf{x} = \mathbf{b}$ exactly.

We solve it by least squares.

Tip: Orthonormal columns should make the problem simple.

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

The normal equations: $\mathbf{Q}^T\mathbf{Q}\hat{\mathbf{x}} = \mathbf{Q}^T\mathbf{b}$

Therefore, $\hat{\mathbf{x}} = \mathbf{Q}^T\mathbf{b}$.

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

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

If \mathbf{Q} has orthonormal columns, the least-squares problem becomes easy:

$\mathbf{Q}\mathbf{x} = \mathbf{b}$	rectangular system ($m > n$) with no solution for most \mathbf{b} .
$\mathbf{Q}^T \mathbf{Q} \hat{\mathbf{x}} = \mathbf{Q}^T \mathbf{b}$	normal equation for the best $\hat{\mathbf{x}}$ — in which $\mathbf{Q}^T \mathbf{Q} = \mathbf{I}$.
$\hat{\mathbf{x}} = \mathbf{Q}^T \mathbf{b}$	\hat{x}_i is $\mathbf{q}_i^T \mathbf{b}$.
$\mathbf{p} = \mathbf{Q} \hat{\mathbf{x}}$	the projection of \mathbf{b} is $(\mathbf{q}_1^T \mathbf{b})\mathbf{q}_1 + (\mathbf{q}_2^T \mathbf{b})\mathbf{q}_2 + \dots + (\mathbf{q}_n^T \mathbf{b})\mathbf{q}_n$.
$\mathbf{p} = \mathbf{Q}\mathbf{Q}^T \mathbf{b}$	the projection matrix is $\mathbf{P} = \mathbf{Q}\mathbf{Q}^T$.

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

Example 4 The following case is simple but typical.

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

63?

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

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

$$\mathbf{q}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \text{ and } (\mathbf{q}_1^T \mathbf{b}) \mathbf{q}_1 = \begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix}; \quad \mathbf{q}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \text{ and } (\mathbf{q}_2^T \mathbf{b}) \mathbf{q}_2 = \begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix}.$$

The overall projection matrix is

$$\mathbf{P} = \mathbf{q}_1 \mathbf{q}_1^T + \mathbf{q}_2 \mathbf{q}_2^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \text{ and } \mathbf{P} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}.$$

Remark: *Projection onto a plane = sum of projections onto orthonormal \mathbf{q}_1 and \mathbf{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 \mathbf{a} , \mathbf{b} , \mathbf{c} orthonormal.
(由 \mathbf{a} , \mathbf{b} , \mathbf{c} 出发, 构造出一组标准正交向量 \mathbf{q}_1 , \mathbf{q}_2 , \mathbf{q}_3 .)

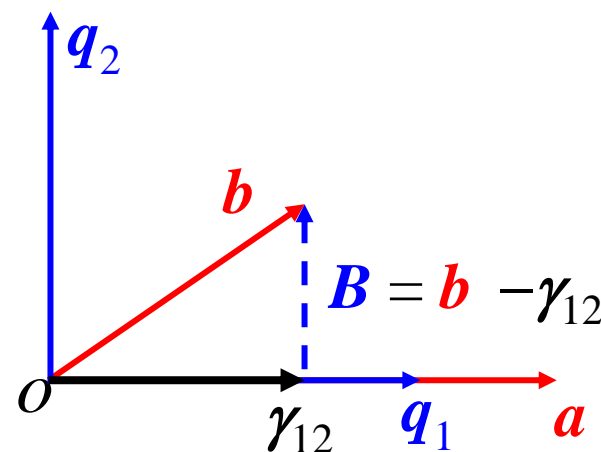
Let $\mathbf{q}_1 = \mathbf{a}/\|\mathbf{a}\|$ to make it a unit vector.

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

$$\gamma_{12} = (\mathbf{q}_1^T \mathbf{b}) \mathbf{q}_1$$

That component has to be subtracted:

Take $\mathbf{B} = \mathbf{b} - \gamma_{12} = \mathbf{b} - (\mathbf{q}_1^T \mathbf{b}) \mathbf{q}_1$,
and $\mathbf{q}_2 = \mathbf{B}/\|\mathbf{B}\|$.



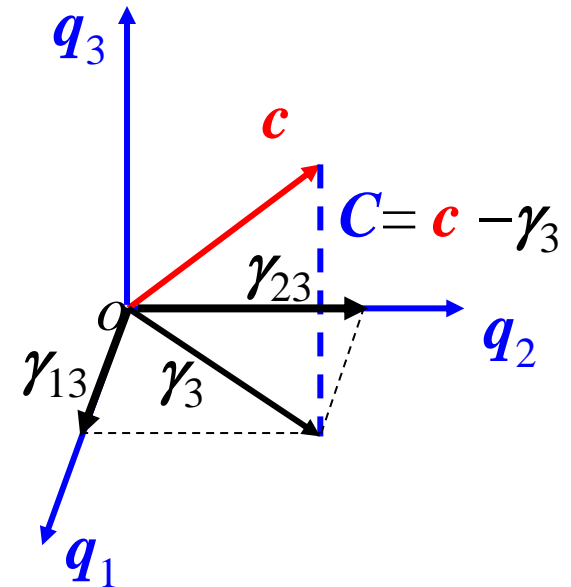
The vector \mathbf{c} will not be in the plane of \mathbf{q}_1 and \mathbf{q}_2 , which is the plane of \mathbf{a} and \mathbf{b} .

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

$$\mathbf{C} = \mathbf{c} - \gamma_3 = \mathbf{c} - (q_1^T \mathbf{c})\mathbf{q}_1 - (q_2^T \mathbf{c})\mathbf{q}_2,$$

and $\mathbf{q}_3 = \mathbf{C}/\|\mathbf{C}\|$.

When there is a fourth vector in \mathbf{R}^n ($n \geq 4$), we subtract away its components in the directions of \mathbf{q}_1 , \mathbf{q}_2 , \mathbf{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 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ and ends with *orthonormal* vectors $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n$.

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

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

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

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

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

Then

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

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

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

Remark:

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

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

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

Remark:

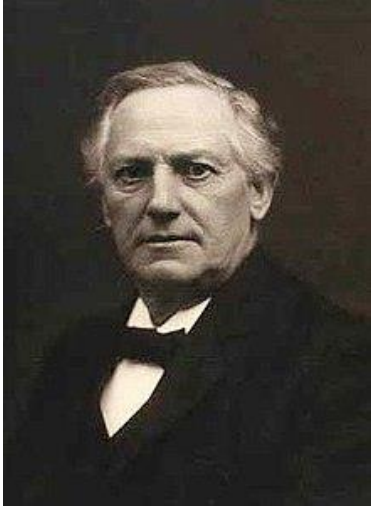
2. It is easier to compute the orthogonal vectors without forcing their lengths to equal one. They can be normalized to unit vector at the end by dividing by their lengths. (先正交化, 再单位化)

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

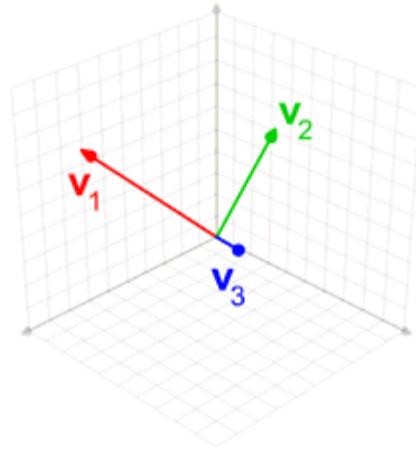
$$\mathbf{B} = \mathbf{b} - \left(\frac{\mathbf{A}^T \mathbf{b}}{\mathbf{A}^T \mathbf{A}} \right) \mathbf{A} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix};$$

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

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



Jørgen Pedersen Gram



Erhard Schmidt

The modified Gram-Schmidt process being executed on three linearly independent, non-orthogonal vectors of a basis for \mathbf{R}^3

https://en.wikipedia.org/wiki/Gram%E2%80%93Schmidt_process

<https://www.geogebra.org/m/stqjb5er>

IV. QR Factorization (QR分解)

The Gram-Schmidt process:

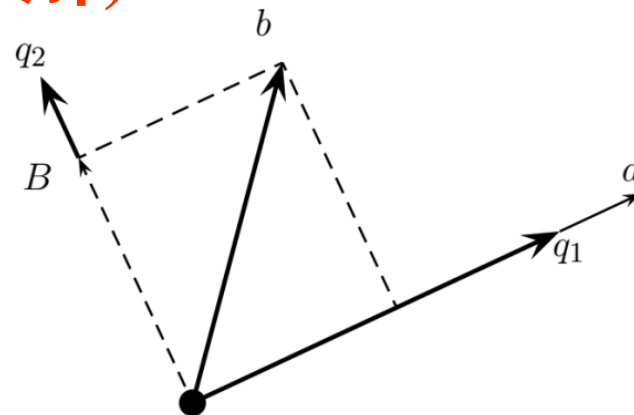
starts with *independent* vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ and ends with *orthonormal* vectors $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n$.

Let $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n]$,

$$\mathbf{Q} = [\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n].$$

The matrices \mathbf{A} and \mathbf{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: \mathbf{R} .



For example,

$$\mathbf{A} = [\mathbf{a}, \mathbf{b}, \mathbf{c}], \quad \mathbf{Q} = [\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3].$$

The idea is to write the $\mathbf{a}, \mathbf{b}, \mathbf{c}$ as combinations of the \mathbf{q} 's:

$$\mathbf{a} = (\mathbf{q}_1^T \mathbf{a}) \mathbf{q}_1,$$

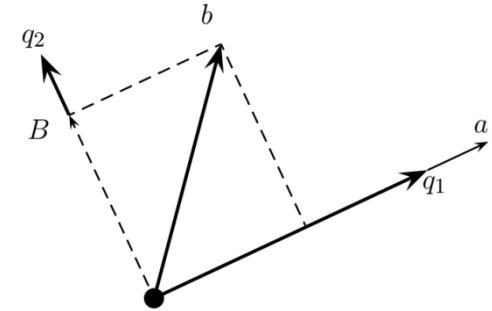
$$\mathbf{b} = (\mathbf{q}_1^T \mathbf{b}) \mathbf{q}_1 + (\mathbf{q}_2^T \mathbf{b}) \mathbf{q}_2,$$

$$\mathbf{c} = (\mathbf{q}_1^T \mathbf{c}) \mathbf{q}_1 + (\mathbf{q}_2^T \mathbf{c}) \mathbf{q}_2 + (\mathbf{q}_3^T \mathbf{c}) \mathbf{q}_3.$$

$$\mathbf{a} = (\mathbf{q}_1^T \mathbf{a}) \mathbf{q}_1,$$

$$\mathbf{b} = (\mathbf{q}_1^T \mathbf{b}) \mathbf{q}_1 + (\mathbf{q}_2^T \mathbf{b}) \mathbf{q}_2,$$

$$\mathbf{c} = (\mathbf{q}_1^T \mathbf{c}) \mathbf{q}_1 + (\mathbf{q}_2^T \mathbf{c}) \mathbf{q}_2 + (\mathbf{q}_3^T \mathbf{c}) \mathbf{q}_3.$$



If we express that in matrix form we have *the new factorization*
 $\mathbf{A} = \mathbf{QR}$:

$$\mathbf{A} = [\mathbf{a} \quad \mathbf{b} \quad \mathbf{c}] = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \mathbf{q}_3] \begin{bmatrix} \mathbf{q}_1^T \mathbf{a} & \mathbf{q}_1^T \mathbf{b} & \mathbf{q}_1^T \mathbf{c} \\ \mathbf{q}_2^T \mathbf{b} & \mathbf{q}_2^T \mathbf{c} \\ \mathbf{q}_3^T \mathbf{c} \end{bmatrix} = \mathbf{QR}.$$

- \mathbf{R} is *upper triangular* because of the way Gram-Schmidt was done. (The first vectors \mathbf{a} and \mathbf{q}_1 fell on the same line. Then \mathbf{q}_1 , \mathbf{q}_2 were in the same plane as \mathbf{a} , \mathbf{b} . The third vectors \mathbf{c} and \mathbf{q}_3 were not involved until step 3.)
- You see the lengths of \mathbf{a} , \mathbf{B} , \mathbf{C} on the diagonal of \mathbf{R} .
- \mathbf{Q} has orthonormal columns.

Example 5 (Continued) Suppose the independent vectors are \mathbf{a} , \mathbf{b} , \mathbf{c} :

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

And we have $\mathbf{Q} = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \mathbf{q}_3] = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{bmatrix}.$

The whole factorization is

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

$$[\mathbf{a} \quad \mathbf{b} \quad \mathbf{c}] = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \mathbf{q}_3] \begin{bmatrix} \mathbf{q}_1^T \mathbf{a} & \mathbf{q}_1^T \mathbf{b} & \mathbf{q}_1^T \mathbf{c} \\ & \mathbf{q}_2^T \mathbf{b} & \mathbf{q}_2^T \mathbf{c} \\ & & \mathbf{q}_3^T \mathbf{c} \end{bmatrix}$$

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

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: $Ax = b$

$A^T A$ becomes easier: $A^T A = R^T Q^T Q R = R^T R$.

The normal equation $A^T A \hat{x} = A^T b$ simplifies to a triangular system:

$$R^T R \hat{x} = R^T Q^T b$$

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

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

