5 Orthogonality and QR

The 'linear algebra way' of talking about "angle" and "similarity" between two vectors is called "inner product". We'll define this next.

So, what is an inner product?

An inner product is a function f of two vector arguments (where both vectors must be from the same vector space V) with the following properties:

(1) $f(\alpha x, y) = \alpha f(x, y)$ for $x, y \in V$

(2) $f(x+y, z) = f(x, z) + f(y, z)$ for $x, y, z \in V$

(3) $f(x, y) = f(y, x)$ for $x, y \in V$

(4) $f(x, x) \geq 0$ for $x \in V$

(5) $f(x, x) = 0 \iff x = 0$ for $x \in V$

(1) and (2) are called "linearity in the first argument"

(3) is called "symmetry"

(4) and (5) are called "positive definiteness"

Can you give an example?

The most important example is the dot product on $\mathbb{R}^n$:

$$f(x,y) = f\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\right) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$= x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

Is any old inner product also "linear" in the second argument?

Check:

$$f(x, y+z) = f(y+z, x) = f(y,x) + f(z,x)$$

$$= f(x,y) + f(x,z) \quad \checkmark$$

$$f(x, \alpha y) = f(\alpha y, x) = \alpha f(y,x) = \alpha f(x,y) \quad \checkmark$$

Do inner products relate to norms at all?

Yes, very much so in fact. Any inner product gives rise to a norm:

$$\|\vec{x}\| := \sqrt{|f(\vec{x},\vec{x})|}$$

For the dot product specifically, we have

$$\sqrt{|\vec{x} \cdot \vec{x}|} = \sqrt{x_1^2 + \cdots + x_n^2} = \|\vec{x}\|_2.$$

Tell me about orthogonality.

Two vectors x and y are called orthogonal if their inner product is 0:

$$f(\vec{x}, \vec{y}) = 0.$$

In this case, the two vectors are also often called perpendicular:

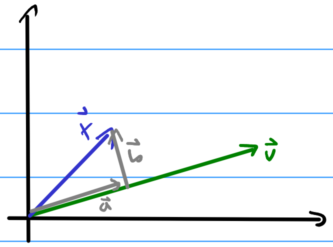
$$x \perp y$$

If f is the dot product, then this means that the two vectors form an angle of 90 degrees.

For orthogonal vectors, we have the Pythagorean theorem:

$$x \perp y \Rightarrow \|x+y\|^2 = \|x\|^2 + \|y\|^2$$

What if I've got two vectors that are not orthogonal, but I'd like them to be?



Given: $\vec{x}, \vec{v}$

Want: $\vec{a}, \vec{b}$

with $\vec{a} = \alpha \vec{v}$

$\vec{b} \perp \vec{v}$

$\vec{x} = \vec{a} + \vec{b}$

also written as: $x^{\parallel v}$ $\vec{x}^{\perp \vec{v}}$

$\rightsquigarrow \quad \vec{x} = \vec{a} + \vec{b} = \alpha \vec{v} + \vec{b}$

$\rightsquigarrow \quad \vec{b} = \vec{x} - \alpha \vec{v}$

$\rightsquigarrow \quad \vec{v} \, (\vec{x} - \alpha \vec{v}) = 0$

$\Leftrightarrow \quad \vec{v} \cdot \vec{x} - \alpha \, \vec{v} \cdot \vec{v} = 0$

$\Leftrightarrow \quad \alpha = \dfrac{\vec{v} \cdot \vec{x}}{\vec{v} \cdot \vec{v}}$

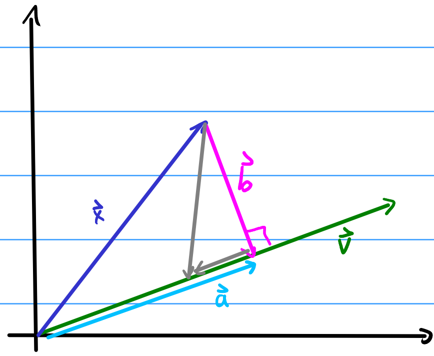
$$\boxed{\begin{aligned} \rightsquigarrow \quad \vec{x}^{\perp \vec{v}} &= \vec{x} - \frac{\vec{v} \cdot \vec{x}}{\vec{v} \cdot \vec{v}} \vec{v} \\ \vec{x}^{\parallel \vec{v}} &= \vec{x} - \vec{x}^{\perp \vec{v}} \end{aligned}}$$

Demo: Orthogonalizing vectors

This process is called orthogonalization.

Note: The expression for α above becomes even simpler if $\|\vec{v}\| = 1$.

In this situation, where is the closest point to $\vec{x}$ on the line $\beta\vec{v}$?

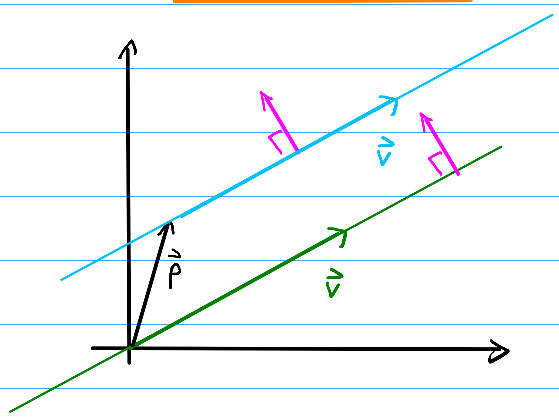


It's the point pointed to by $\vec{a}$.

By the Pythagorean theorem, all other points on the line would have a larger norm:

They would contain $\vec{b}$ and a component along $\vec{v}$.

How does the point-normal form (of a line) work?



1. Find a vector $\vec{b}$ orthogonal to the vector $\vec{v}$ spanning the line. (For example by orthogonalization of another vector against $\vec{v}$.)

2. Scale $\vec{n}$ to have norm 1: $\vec{n} = \frac{\vec{b}}{\|\vec{b}\|}$.
$\vec{n}$ is called the (unit) normal.

3. Find a point $\vec{p}$ on the line.

4. Compute $r = \vec{p} \cdot \vec{n}$

5. The equation $(\vec{x} - \vec{p}) \cdot \vec{n} = \vec{x} \cdot \vec{n} - r = 0$
is satisfied exactly by points $\vec{x}$ on the line.

6. For all other points, $\vec{x} \cdot \vec{n} - r$ gives the (signed) distance from the line.

Demo: Point-normal form and signed distance

What's an orthogonal basis?

A basis $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n$ of a vector space that is also pairwise orthogonal:

$$b_i \perp b_j \quad \text{if} \quad i \neq j$$

What's an orthonormal basis (or "ONB")?

An orthogonal basis where each basis vector $\vec{b}_i$ has norm 1.

For some given vector $\vec{x}$, how do I find coefficients with respect to an ONB?

$$\vec{x} = \underbrace{(x \cdot b_1)}\vec{b}_1 + \underbrace{(x \cdot b_2)}\vec{b}_2 + \underbrace{(x \cdot b_3)}\vec{b}_3 + \cdots + \underbrace{(x \cdot b_n)}\vec{b}_n$$

Much easier than finding coefficients by solving a linear system!

Also much cheaper: $O(n^2)$

Can we build a matrix that computes those coefficients for us?

$$Q = \begin{pmatrix} | & & | \\ \vec{b}_1 & \cdots & \vec{b}_n \\ | & & | \end{pmatrix} \rightsquigarrow Q^T\vec{x} = \begin{pmatrix} \vec{b}_1 \cdot \vec{x} \\ \vdots \\ \vec{b}_n \cdot \vec{x} \end{pmatrix}$$

A square matrix whose columns are orthonormal is called orthogonal.

⚠

What else is true for a orthogonal matrix Q?

$$Q^TQ = I \quad \rightsquigarrow \quad Q^{-1} = Q^T \quad \rightsquigarrow \quad QQ^T = I$$

$$\|Qx\|_2^2 = (Qx)\cdot(Qx) = (Qx)^T(Qx) = x^TQ^TQx = x^Tx = \|x\|_2^2 \; !$$

In words: Qx has the same 2-norm as x.

What if Q contains a few zero columns instead of orthonormal vectors?

$$Q = \left[\begin{array}{ccc|c} | & & | & \\ \vec{b}_1 & \cdots & \vec{b}_k & 0 \\ | & & | & \end{array}\right]$$

orthonormal columns

Define $P := QQ^T$

Compute $P\vec{x}$ for $\vec{x} = \alpha_1\vec{b}_1 + \cdots + \alpha_n\vec{b}_n$:

$$Q^T\vec{x} = \begin{pmatrix} \vec{x}\cdot\vec{b}_1 \\ \vdots \\ \vec{x}\cdot\vec{b}_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$Q(Q^T\vec{x}) = \underbrace{(\vec{x}\cdot\vec{b}_1)\vec{b}_1 + \cdots + (\vec{x}\cdot\vec{b}_k)\vec{b}_k}_{x \text{ projected onto } \vec{b}_1 \ldots \vec{b}_k} + \underbrace{0\cdot\vec{b}_{k+1} + \cdots + 0\cdot\vec{b}_n}_{\vec{0}}$$

P is called the orthogonal projector onto $\vec{b}_1 \ldots \vec{b}_k$.

Observe: $P^2 = P$. Demo: Orthogonal Projection

Orthonormal vectors seem very useful. How can we make more than two?

Need pairwise orthogonality!

$\cdot$	$\vec{x}_1$	$\vec{x}_2$	$\vec{x}_3$
$\vec{x}_1$	1	0	0
$\vec{x}_2$	0	1	0
$\vec{x}_3$	0	0	1

needs to be orthogonalized against all prior vectors!

Demo: Orthogonalizing three vectors

Demo: Gram-Schmidt--The Movie

Demo: Gram-Schmidt and Modified Gram-Schmidt

Demo: Keeping track of the coefficients in Gram-Schmidt
(QR factorization)

So, what is the QR factorization?

The QR factorization is given by

$$A = QR$$

where A is any matrix,
Q is orthogonal, and
R is upper triangular.

If life were consistent, shouldn't this be called the QU factorization?

Yes.

What is the cost of QR factorization (for an nxn matrix A)?

For each of n columns of A:

Orthogonalize against every one of (at most n) previous vectors:

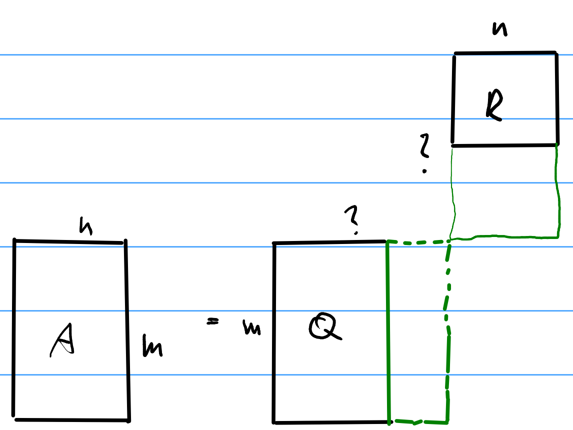
By computing a dot product at a cost of O(n).

$\rightarrow O(n^3)$

Demo: Complexity of LU and QR

Does QR work for non-square matrices?

This is very similar to LU factorization.



A: $m \times n$, $m \geq n$

Either:

Q $m \times m$ square
R: $m \times n$
"full" "complete"

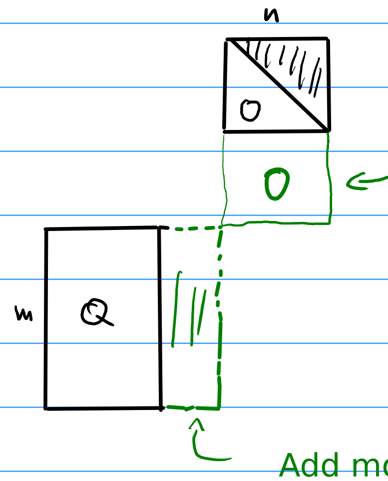
Q: $m \times n$
R: $n \times n$ square
"thin" "reduced"

Is Q still orthogonal in a "thin" QR factorization?

No--its columns do not form a full basis of $\mathbb{R}^n$.

That's why it is sometimes reasonable to ask for a "full"/"complete" QR.

If I have a "thin" QR, can I obtain a "full" QR from it?



Add zeros here.

Add more orthogonal columns here.

How?

(1) Randomly choose m-n vectors.

(2) Use Gram-Schmidt to orthogonalize them against all previous columns.

Can QR fail?

If two columns are exactly linearly dependent, we may obtain a zero vector when orthogonalizing.

-> Zero on diagonal in R.

Not much of an issue: Does not affect existence of QR.

Pivoted QR exists, but is not very common.