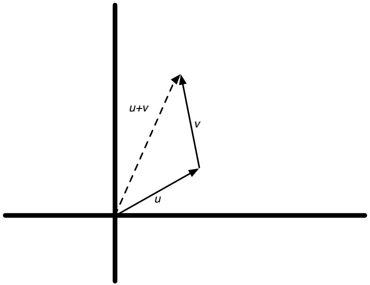
Orthogonality

Orthogonal is linear-algebra-ese for *perpendicular*.
We'll define it so as to make Pythagorean Theorem true.
Let $\mathbf{u}$ and $\mathbf{v}$ be vectors.
Their lengths are $||\mathbf{u}||$ and $||\mathbf{v}||$.
Draw the corresponding arrows, and the arrow for $\mathbf{u}+\mathbf{v}$
The arrow for $\mathbf{u}+\mathbf{v}$ is the "hypotenuse". (The triangle is not necessarily a right angle.)



The squared length of the vector $\mathbf{u}+\mathbf{v}$ (the "hypotenuse") is

$$\begin{aligned} ||\mathbf{u}+\mathbf{v}||^2 &= \langle \mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{v} \rangle \\ &= \langle \mathbf{u}, \mathbf{u}+\mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{u}+\mathbf{v} \rangle && \text{by linearity of inner product in } 1^{st} \text{ argument} \\ &= \langle \mathbf{u}, \mathbf{u} \rangle + \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle && \text{by symmetry and linearity} \\ &= ||\mathbf{u}||^2 + 2\langle \mathbf{u}, \mathbf{v} \rangle + ||\mathbf{v}||^2 && \text{by symmetry} \end{aligned}$$

The last expression is $||\mathbf{u}||^2 + ||\mathbf{v}||^2$ if and only if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.
We therefore define $\mathbf{u}$ and $\mathbf{v}$ to be *orthogonal* if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.
Pythagorean Theorem for vectors: if vectors $\mathbf{u}$ and $\mathbf{v}$ over the reals are orthogonal then $||\mathbf{u}+\mathbf{v}||^2 = ||\mathbf{u}||^2 + ||\mathbf{v}||^2$

Properties of orthogonality

To solve the Fire Engine Problem, we will use the Pythagorean Theorem in conjunction with the following simple observations:

Orthogonality Properties:

Property O1: If $\mathbf{u}$ is orthogonal to $\mathbf{v}$ then $\mathbf{u}$ is orthogonal to $\alpha\,\mathbf{v}$ for every scalar α.

Property O2: If $\mathbf{u}$ and $\mathbf{v}$ are both orthogonal to $\mathbf{w}$ then $\mathbf{u}+\mathbf{v}$ is orthogonal to $\mathbf{w}$.

Proof:

1. $\langle \mathbf{u}, \alpha\,\mathbf{v}\rangle = \alpha\,\langle \mathbf{u}, \mathbf{v}\rangle = \alpha\,0 = 0$
2. $\langle \mathbf{u}+\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{u}, \mathbf{w}\rangle + \langle \mathbf{v}, \mathbf{w}\rangle = 0 + 0$

Example: $[1,2]\cdot[2,-1] = 0$ so $[1,2]\cdot[20,-10] = 0$

Example:

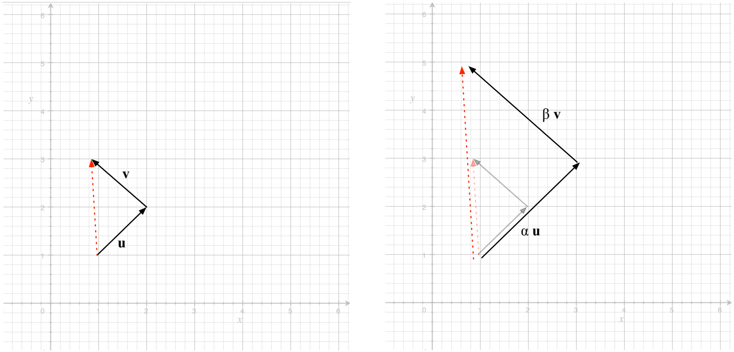
$$\begin{array}{rcl} [1,2,1]\cdot[1,-1,1] & = & 0 \\ [0,1,1]\cdot[1,-1,1] & = & 0 \\ \hline ([1,2,1]+[0,1,1])\cdot[1,-1,1] & = & 0 \end{array}$$

Length of sum of orthogonal vectors

Scaling orthogonal vectors gives us orthogonal vectors.

Lemma: If **u** is orthogonal to **v** then, for any scalars α, β,

$$||\alpha\,\mathbf{u} + \beta\,\mathbf{v}||^2 = \alpha^2||\mathbf{u}||^2 + \beta^2||\mathbf{v}||^2$$



Length of sum of orthogonal vectors

Scaling orthogonal vectors gives us orthogonal vectors.

Lemma: If $\mathbf{u}$ is orthogonal to $\mathbf{v}$ then, for any scalars α, β,

$$||\alpha\,\mathbf{u} + \beta\,\mathbf{v}||^2 = \alpha^2||\mathbf{u}||^2 + \beta^2||\mathbf{v}||^2$$

Proof:

$$\begin{aligned}
(\alpha\,\mathbf{u} + \beta\,\mathbf{v}) \cdot (\alpha\,\mathbf{u} + \beta\,\mathbf{v}) &= \alpha\,\mathbf{u} \cdot \alpha\,\mathbf{u} + \beta\,\mathbf{v} \cdot \beta\,\mathbf{v} + \alpha\,\mathbf{u} \cdot \beta\,\mathbf{v} + \beta\,\mathbf{v} \cdot \alpha\,\mathbf{u} \\
&= \alpha\,\mathbf{u} \cdot \alpha\,\mathbf{u} + \beta\,\mathbf{v} \cdot \beta\,\mathbf{v} + \alpha\beta\,(\mathbf{u} \cdot \mathbf{v}) + \beta\alpha\,(\mathbf{v} \cdot \mathbf{u}) \\
&= \alpha\,\mathbf{u} \cdot \alpha\,\mathbf{u} + \beta\,\mathbf{v} \cdot \beta\,\mathbf{v} + 0 + 0 \\
&= \alpha^2\,\|\mathbf{u}\|^2 + \beta^2\,\|\mathbf{v}\|^2
\end{aligned}$$

Mutual orthogonality

Definition: We say $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are *mutually orthogonal* if $\mathbf{v}_i$ is orthogonal to $\mathbf{v}_j$ for every pair i, j such that $i \neq j$.

Example: $[1,2,1], [1,-1,1], [1,0,-1]$ are mutually orthogonal:

- $\langle [1,2,1], [1,-1,1] \rangle = 0$
- $\langle [1,2,1], [1,0,-1] \rangle = 0$
- $\langle [1,-1,1], [1,0,-1] \rangle = 0$

Lemma: If $\mathbf{u}$ is orthogonal to $\mathbf{v}$ then, for any scalars α, β,

$$||\alpha\, \mathbf{u} + \beta\, \mathbf{v}||^2 = \alpha_1^2 ||\mathbf{u}||^2 + \beta^2 ||\mathbf{v}||^2$$

generalizes to:

Lemma: If $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are mutually orthogonal then, for any scalars $\alpha_1, \ldots, \alpha_n$,

$$||\alpha_1 \mathbf{v}_1 + \cdots + \alpha_n \mathbf{v}_n||^2 = \alpha_1^2 ||\mathbf{v}_1||^2 + \cdots + \alpha_n^2 ||\mathbf{v}_n||^2$$

Orthogonality helps solve the *fire engine* problem

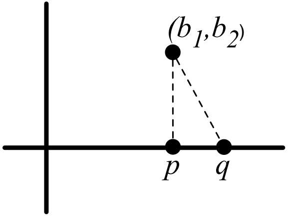
Fire Engine Lemma:

- Let $\mathbf{b}$ be a vector.
- Let $\mathbf{a}$ be a nonzero vector $\Rightarrow$ The set $\{\alpha\,\mathbf{a}\ :\ \alpha \in \mathbb{R}\}$ is a line L
- Let $\mathbf{p}$ be the point on the line L such that $\mathbf{b} - \mathbf{p}$ is orthogonal to $\mathbf{a}$.

Then $\mathbf{p}$ is the point on the line that is closest to $\mathbf{b}$.

Example: Line is the x-axis, i.e. the set $\{(x, y) :\ \ y = 0\}$, and point is (b_1, b_2).

Lemma states: closest point on the line is $\mathbf{p} = (b_1, 0)$.



- For any other point $\mathbf{q}$, the points (b_1, b_2), $\mathbf{p}$, and $\mathbf{q}$ form a right triangle.
- Since $\mathbf{q}$ is different from $\mathbf{p}$, the base is nonzero.
- By the Pythagorean Theorem, the hypotenuse's length is greater than the height.
- This shows that $\mathbf{q}$ is farther from (b_1, b_2) than $\mathbf{p}$ is.

Orthogonality helps solve the *fire engine* problem

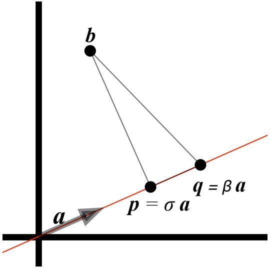
Fire Engine Lemma:

- Let $\mathbf{b}$ be a vector.
- Let $\mathbf{a}$ be a nonzero vector $\Rightarrow$ The set $\{\alpha\,\mathbf{a} \;:\; \alpha \in \mathbb{R}\}$ is a line L
- Let $\mathbf{p}$ be the point on the line L such that $\mathbf{b} - \mathbf{p}$ is orthogonal to $\mathbf{a}$.

Then $\mathbf{p}$ is the point on the line that is closest to $\mathbf{b}$.

Proof: Let $\mathbf{q}$ be any point on L. The three points $\mathbf{q}$, $\mathbf{p}$, and $\mathbf{b}$ form a triangle.

- Since $\mathbf{p}$ and $\mathbf{q}$ are both on L, they are both multiples of $\mathbf{a}$, so their difference $\mathbf{p} - \mathbf{q}$ is also a multiple of $\mathbf{a}$.
- Since $\mathbf{b} - \mathbf{p}$ is orthogonal to $\mathbf{a}$, therefore, it is also orthogonal to $\mathbf{p} - \mathbf{q}$



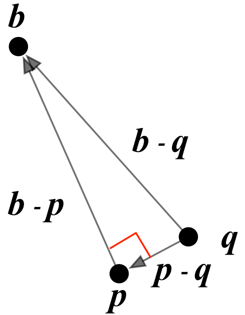
Orthogonality helps solve the *fire engine* problem

Fire Engine Lemma:

- Let $\mathbf{b}$ be a vector.
- Let $\mathbf{a}$ be a nonzero vector $\Rightarrow$ The set $\{\alpha\,\mathbf{a} \; : \; \alpha \in \mathbb{R}\}$ is a line L
- Let $\mathbf{p}$ be the point on the line L such that $\mathbf{b} - \mathbf{p}$ is orthogonal to $\mathbf{a}$.

Then $\mathbf{p}$ is the point on the line that is closest to $\mathbf{b}$.

Proof: Let $\mathbf{q}$ be any point on L. The three points $\mathbf{q}$, $\mathbf{p}$, and $\mathbf{b}$ form a triangle.



- Since $\mathbf{p}$ and $\mathbf{q}$ are both on L, they are both multiples of $\mathbf{a}$, so their difference $\mathbf{p} - \mathbf{q}$ is also a multiple of $\mathbf{a}$.
- Since $\mathbf{b} - \mathbf{p}$ is orthogonal to $\mathbf{a}$, therefore, it is also orthogonal to $\mathbf{p} - \mathbf{q}$
- Hence by the Pythagorean Theorem,

$$||\mathbf{b} - \mathbf{q}||^2 = ||\mathbf{p} - \mathbf{q}||^2 + ||\mathbf{b} - \mathbf{p}||^2$$

- If $\mathbf{q} \neq \mathbf{p}$ then $||\mathbf{p} - \mathbf{q}||^2 > 0$ so $||\mathbf{b} - \mathbf{p}|| < ||\mathbf{b} - \mathbf{q}||$.

Decomposition of $\mathbf{b}$ into parallel and perpendicular components

Lemma states: among all the points on the line $\{\alpha\,\mathbf{a} \;:\; \alpha \in R\}$, the closest to $\mathbf{b}$ is the point $\mathbf{p}$ on such that $\mathbf{b} - \mathbf{p}$ is orthogonal to $\mathbf{a}$.

Definition: For any vector $\mathbf{b}$ and any vector $\mathbf{a}$, define vectors $\mathbf{b}^{||\mathbf{a}}$ and $\mathbf{b}^{\perp\mathbf{a}}$ to be the *projection of* $\mathbf{b}$ *onto Span* $\{\mathbf{a}\}$ and the *projection of* $\mathbf{b}$ *orthogonal to* $\mathbf{a}$ if

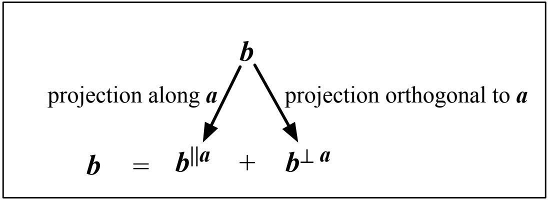
$$\mathbf{b} = \mathbf{b}^{||\mathbf{a}} + \mathbf{b}^{\perp\mathbf{a}}$$

and there is a scalar $\sigma \in R$ such that

$$\mathbf{b}^{||\mathbf{a}} = \sigma\,\mathbf{a}$$

and

$$\mathbf{b}^{\perp\mathbf{a}} \text{ is orthogonal to } \mathbf{a}$$



Closest-Point Corollary

For any vector $\mathbf{b}$ and any vector $\mathbf{a}$, define vectors $\mathbf{b}^{||\mathbf{a}}$ and $\mathbf{b}^{\perp\mathbf{a}}$ ....

$$\mathbf{b} = \mathbf{b}^{||\mathbf{a}} + \mathbf{b}^{\perp\mathbf{a}}$$

and there is a scalar $\sigma \in R$ such that

$$\mathbf{b}^{||\mathbf{a}} = \sigma\,\mathbf{a}$$

and

$$\mathbf{b}^{\perp\mathbf{a}} \text{ is orthogonal to } \mathbf{a}$$

Closest-Point Corollary: For any vector $\mathbf{b}$ and vector $\mathbf{a}$ over the reals,

- the point in Span $\{\mathbf{a}\}$ that is closest to $\mathbf{b}$ is the projection $\mathbf{b}^{||\mathbf{a}}$ onto Span $\{\mathbf{a}\}$,
- and the distance between that point and $\mathbf{b}$ is $\|\mathbf{b}^{\perp\mathbf{a}}\|$, the norm of the projection of $\mathbf{b}$ orthogonal to $\mathbf{a}$.

Decomposition of $\mathbf{b}$ into parallel and perpendicular components: example

For any vector $\mathbf{b}$ and any vector $\mathbf{a}$, define vectors $\mathbf{b}^{||\mathbf{a}}$ and $\mathbf{b}^{\perp\mathbf{a}}$ ....

$$\mathbf{b} = \mathbf{b}^{||\mathbf{a}} + \mathbf{b}^{\perp\mathbf{a}}$$

and there is a scalar $\sigma \in R$ such that

$$\mathbf{b}^{||\mathbf{a}} = \sigma\,\mathbf{a}$$

and

$$\mathbf{b}^{\perp\mathbf{a}} \text{ is orthogonal to } \mathbf{a}$$

Example: What if $\mathbf{a}$ is the zero vector?
In this case, the only vector $\mathbf{b}^{||\mathbf{a}}$ satisfying the second equation is the zero vector.
According to first equation, $\mathbf{b}^{\perp\mathbf{a}}$ must equal $\mathbf{b}$.
Fortunately, this choice of $\mathbf{b}^{\perp\mathbf{a}}$ does satisfy third equation: $\mathbf{b}^{\perp\mathbf{a}}$ is orthogonal to $\mathbf{a}$.
Indeed, *every* vector is orthogonal to $\mathbf{a}$ when $\mathbf{a}$ is the zero vector.
What is the point in Span $\{\mathbf{0}\}$ closest to $\mathbf{b}$?
The *only* point in Span $\{\mathbf{0}\}$ is the zero vector...
so that must be the closest point to $\mathbf{b}$, and the distance to $\mathbf{b}$ is $||\mathbf{b}||$.

Computing the projections

$\mathbf{b} = \mathbf{b}^{||\mathbf{a}} + \mathbf{b}^{\perp\mathbf{a}}$ $\mathbf{b}^{||\mathbf{a}} = \sigma\,\mathbf{a}$ $\mathbf{b}^{\perp\mathbf{a}}$ is orthogonal to $\mathbf{a}$

If $\mathbf{a} = \mathbf{0}$ then $\mathbf{b}^{||\mathbf{a}} = \mathbf{0}$

What if $\mathbf{a} \neq \mathbf{0}$? Need to compute σ....

- $\left\langle \mathbf{b}^{\perp\mathbf{a}}, \mathbf{a} \right\rangle = 0$. Substitute for $\mathbf{b}^{\perp\mathbf{a}}$: $\left\langle \mathbf{b} - \mathbf{b}^{||\mathbf{a}}, \mathbf{a} \right\rangle = 0$.
- Substitute for $\mathbf{b}^{||}$: $\langle \mathbf{b} - \sigma\,\mathbf{a}, \mathbf{a} \rangle = 0$.
- Using linearity and homogeneity of inner product,

$$\langle \mathbf{b}, \mathbf{a} \rangle - \sigma\, \langle \mathbf{a}, \mathbf{a} \rangle = 0$$

- Solving for σ, we obtain

$$\sigma = \frac{\langle \mathbf{b}, \mathbf{a} \rangle}{\langle \mathbf{a}, \mathbf{a} \rangle}$$

In the special case in which $\|\mathbf{a}\| = 1$, the denominator $\langle \mathbf{a}, \mathbf{a} \rangle = 1$ so

$$\sigma = \langle \mathbf{b}, \mathbf{a} \rangle$$

Quiz: Write `project_along(b, a)` to return the vector $\mathbf{b}^{||\mathbf{a}}$

Computing the projections

$\mathbf{b} = \mathbf{b}^{||\mathbf{a}} + \mathbf{b}^{\perp\mathbf{a}}$ $\quad$ $\mathbf{b}^{||\mathbf{a}} = \sigma\,\mathbf{a}$ $\quad$ $\mathbf{b}^{\perp\mathbf{a}}$ is orthogonal to $\mathbf{a}$

If $\mathbf{a} = \mathbf{0}$ then $\mathbf{b}^{||\mathbf{a}} = \mathbf{0}$

What if $\mathbf{a} \neq \mathbf{0}$? Need to compute σ....

- $\left\langle \mathbf{b}^{\perp\mathbf{a}}, \mathbf{a} \right\rangle = 0$. Substitute for $\mathbf{b}^{\perp\mathbf{a}}$: $\left\langle \mathbf{b} - \mathbf{b}^{||\mathbf{a}}, \mathbf{a} \right\rangle = 0$.
- Substitute for $\mathbf{b}^{||}$: $\langle \mathbf{b} - \sigma\,\mathbf{a}, \mathbf{a} \rangle = 0$.
- Using linearity and homogeneity of inner product,

$$\langle \mathbf{b}, \mathbf{a} \rangle - \sigma\,\langle \mathbf{a}, \mathbf{a} \rangle = 0$$

- Solving for σ, we obtain

$$\sigma = \frac{\langle \mathbf{b}, \mathbf{a} \rangle}{\langle \mathbf{a}, \mathbf{a} \rangle}$$

In the special case in which $\|\mathbf{a}\| = 1$, the denominator $\langle \mathbf{a}, \mathbf{a} \rangle = 1$ so

$$\sigma = \langle \mathbf{b}, \mathbf{a} \rangle$$

Quiz: Write `project_along(b, a)` to return the vector $\mathbf{b}^{||\mathbf{a}}$

Answer: `def project_along(b, a):  return ((b*a)/(a*a))*a` Almost.

Computing the projections

$\mathbf{b} = \mathbf{b}^{||\mathbf{a}} + \mathbf{b}^{\perp\mathbf{a}}$ $\quad$ $\mathbf{b}^{||\mathbf{a}} = \sigma\, \mathbf{a}$ $\quad$ $\mathbf{b}^{\perp\mathbf{a}}$ is orthogonal to $\mathbf{a}$

If $\mathbf{a} = \mathbf{0}$ then $\mathbf{b}^{||\mathbf{a}} = \mathbf{0}$

What if $\mathbf{a} \neq \mathbf{0}$? Need to compute σ....

- $\left\langle \mathbf{b}^{\perp\mathbf{a}}, \mathbf{a} \right\rangle = 0$. Substitute for $\mathbf{b}^{\perp\mathbf{a}}$: $\left\langle \mathbf{b} - \mathbf{b}^{||\mathbf{a}}, \mathbf{a} \right\rangle = 0$.
- Substitute for $\mathbf{b}^{||}$: $\langle \mathbf{b} - \sigma\, \mathbf{a}, \mathbf{a} \rangle = 0$.
- Using linearity and homogeneity of inner product,

$$\langle \mathbf{b}, \mathbf{a} \rangle - \sigma\, \langle \mathbf{a}, \mathbf{a} \rangle = 0$$

- Solving for σ, we obtain

$$\sigma = \frac{\langle \mathbf{b}, \mathbf{a} \rangle}{\langle \mathbf{a}, \mathbf{a} \rangle}$$

In the special case in which $\|\mathbf{a}\| = 1$, the denominator $\langle \mathbf{a}, \mathbf{a} \rangle = 1$ so

$$\sigma = \langle \mathbf{b}, \mathbf{a} \rangle$$

Quiz: Write `project_along(b, a)` to return the vector $\mathbf{b}^{||\mathbf{a}}$

Answer: `def project_along(b, a): return ((b*a)/(a*a))*a` Almost.

Best:

```
def project_along(b, a): return ((b*a)/(a*a) if a*a != 0 else 0)*a
```

Computing the projections

$\mathbf{b} = \mathbf{b}^{||\mathbf{a}} + \mathbf{b}^{\perp\mathbf{a}}$ $\mathbf{b}^{||\mathbf{a}} = \sigma\,\mathbf{a}$ $\mathbf{b}^{\perp\mathbf{a}}$ is orthogonal to $\mathbf{a}$

- $\sigma = \dfrac{\langle \mathbf{b}, \mathbf{a} \rangle}{\langle \mathbf{a}, \mathbf{a} \rangle}$
- However, if $\mathbf{a} = \mathbf{0}$ then $\sigma = 0$.
-
```
def project_along(b, a):
    sigma = (b*a)/(a*a) if a*a != 0 else 0
    return sigma * a
```

Quiz: Use `project_along(b, a)` to write the procedure

`project_orthogonal_1(b, a)`

that returns $\mathbf{b}^{\perp\mathbf{a}}$

Computing the projections

$\mathbf{b} = \mathbf{b}^{||\mathbf{a}} + \mathbf{b}^{\perp\mathbf{a}}$ $\mathbf{b}^{||\mathbf{a}} = \sigma\,\mathbf{a}$ $\mathbf{b}^{\perp\mathbf{a}}$ is orthogonal to $\mathbf{a}$

- $\sigma = \dfrac{\langle \mathbf{b},\mathbf{a}\rangle}{\langle \mathbf{a},\mathbf{a}\rangle}$
- However, if $\mathbf{a} = \mathbf{0}$ then $\sigma = 0$.
- ```
  def project_along(b, a):
   sigma = (b*a)/(a*a) if a*a != 0 else 0
   return sigma * a
  ```

Quiz: Use `project_along(b, a)` to write the procedure

`project_orthogonal_1(b, a)`

that returns $\mathbf{b}^{\perp\mathbf{a}}$

```
def project_orthogonal_1(b, a):  return b - project_along(b, a)
```
```

Projecting along “nearly zero” vectors

Mathematically, this procedure is correct:

```
def project_along(b, a):
 sigma = (b*a)/(a*a) if a*a != 0 else 0
 return sigma * a
```

However, because of floating-point roundoff error, we need to make a slight change.

Often the vector **a** will be not a truly zero vector but practically it will be zero.

If the entries of **a** are tiny, the procedure should treat **a** as a zero vector: `sigma` should be assigned zero.

We will consider **a** to be a zero vector if its squared norm is no more than, say, 10^{-20}.

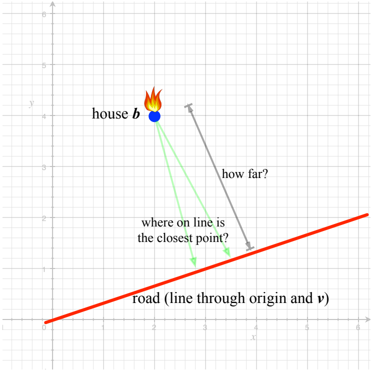
Revised version:

```
def project_along(b, a):
 sigma = (b*a)/(a*a) if a*a > 1e-20 else 0
 return sigma * a
```

Solution to the *fire engine* problem

Example:
$\mathbf{a} = [6, 2]$ and $\mathbf{b} = [2, 4]$.

The closest point on the line $\{\alpha\, \mathbf{a} \;:\; \alpha \in \mathbb{R}\}$ is the point $\mathbf{b}^{||\mathbf{a}} = \sigma\, \mathbf{a}$ where



$$\begin{aligned}
\sigma &= \frac{\mathbf{a} \cdot \mathbf{b}}{\mathbf{a} \cdot \mathbf{a}} \\
&= \frac{6 \cdot 2 + 2 \cdot 4}{6 \cdot 6 + 2 \cdot 2} \\
&= \frac{20}{40} \\
&= \frac{1}{2}
\end{aligned}$$

Thus the point closest to $\mathbf{b}$ is $\frac{1}{2}\,[6, 2] = [3, 1]$.
The distance to $\mathbf{b}$ is

$$\|\mathbf{b}^{\perp \mathbf{a}}\| = \|[2, 4] - [3, 1]\| = \|[-1, 3]\| = \sqrt{10}$$

which is just under 3.5, the length of the firehose.
The house is saved!

Best approximation

The *fire engine* problem can be restated as finding the vector on the line that "best approximates" the given vector $\mathbf{b}$.

By "best approximation", we just mean closest.

This notion of "best approximates" comes up again and again:

- in least-squares, a fundamental data analysis technique,
- image compression,
- in principal component analysis, another data analysis technique, and
- in latent semantic analysis, an information retrieval technique.

Towards solving the higher-dimensional version of *best approximation*

The fire engine problem can be stated thus:

> **Computational Problem:** *Closest point in the span of a single vector*
> Given a vector $\mathbf{b}$ and a vector $\mathbf{a}$ over the reals, find the vector in Span $\{\mathbf{a}\}$ closest to $\mathbf{b}$.

A natural generalization of the *fire engine* problem is this:

> **Computational Problem:** *Closest point in the span of several vectors*
> Given a vector $\mathbf{b}$ and vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ over the reals, find the vector in Span $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ closest to $\mathbf{b}$.

We will study this problem next.