Quiz

You are given a procedure `transformation(A)` with the following spec:

- *input:* Mat A
- *output:* Mat M such that $M * A$ is a Mat in echelon form

Your job is to write each of the following procedures:

```
A = Mat(({'a','b'},{'A','B'}), {('a','A'):one,
  ('b','B'):one,('a','B'):one, ('b','A'):one})
>>> M=transformation(A)
>>> print(M*A)

          A    B
      ---------
 0  |   one  one
 1  |     0    0
```

The input to these is a list of Vecs:

- `rank(L)`
- `is_independent(L)`
- `basis(L)` returns a list of Vecs forming a basis for Span L
- `null_space_basis(A)` where `A` is a `Mat`
- `solve(A, b)` where `A` is a `Mat` and `b` is a `Vec`

For `solve`, you can also assume a procedure `echelon_solve(A,b)` that requires A to be in echelon form and solves $A\mathbf{x} = \mathbf{b}$.

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:

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

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

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

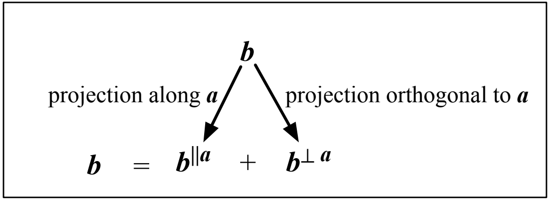
$$\mathbf{b} = \mathbf{b}^{||\mathbf{a}} + \mathbf{b}^{\perp\mathbf{a}}$$

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

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

and

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



Decomposition of $\mathbf{b}$ into $\mathbf{b}^{||\mathbf{a}}$ 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 $\mathbf{b}^{\perp\mathbf{a}}$ is orthogonal to $\mathbf{a}$

Example: $\mathbf{b} = [b_1, b_2]$, $\mathbf{a} = [1, 0]$.
Then $\mathbf{b}^{||\mathbf{a}} = [b_1, 0]$ Note that $[b_1, 0] = \mathbf{b}_1\ [1, 0]$.

$$\mathbf{b}^{\perp\mathbf{a}} = \mathbf{b} - \mathbf{b}^{||\mathbf{a}} = [b_1, b_2] - [b_1, 0] = [0, b_2]$$

Example: $\mathbf{b} = [10, 20, 30]$ and $\mathbf{a} = [-1, 2, 1]$.
I claim $\mathbf{b}^{||\mathbf{a}} = [-10, 20, 10]$ and therefore $\mathbf{b}^{\perp\mathbf{a}} = [10, 20, 30] - [-10, 20, 10] = [20, 0, 20]$.

Are these correct?

- Check if $\mathbf{b}^{||\mathbf{a}} = \sigma\mathbf{a}$ for some σ... Yes, $\sigma = 10$
- Check if $\mathbf{b}^{\perp\mathbf{a}}$ is orthogonal to $\mathbf{a}$... $[20, 0, 20] \cdot [-1, 2, 1] = 0$, so yes

Orthogonality helps solve the *fire engine* problem

Fire Engine Lemma:

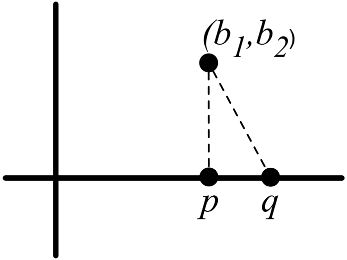
- Let $\mathbf{b}$ be a vector.
- Let $\mathbf{a}$ be a nonzero vector
-

Then $\mathbf{b}^{||\mathbf{a}}$ is the point on the line $\{\alpha\, \mathbf{a}\ :\ \alpha \in \mathbb{R}\}$ that is closest to $\mathbf{b}$, and the distance is $\|\mathbf{b}^{\perp\mathbf{a}}\|$.

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 $\mathbf{b} = (b_1, b_2)$, $\mathbf{b}^{||\mathbf{a}}$, and $\mathbf{q}$ form a right triangle.
- Since $\mathbf{q}$ is different from $\mathbf{b}^{||\mathbf{a}}$, the base is nonzero.
- By the Pythagorean Theorem, the hypotenuse's length is greater than the height.



Orthogonality helps solve the *fire engine* problem

Fire Engine Lemma:

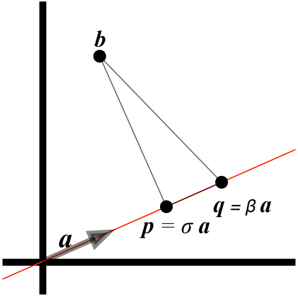
- Let $\mathbf{b}$ be a vector.
- Let $\mathbf{a}$ be a nonzero vector
-

Then $\mathbf{b}^{||\mathbf{a}}$ is the point on the line $\{\alpha\,\mathbf{a}\ :\ \alpha \in \mathbb{R}\}$ that is closest to $\mathbf{b}$, and the distance is $||\mathbf{b}^{\perp\mathbf{a}}||$.

Proof: Let L be the line. Let $\mathbf{p} = \mathbf{b}^{||\mathbf{a}}$. 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 $\mathbf{a}$.
- Hence, since $\mathbf{b}^{\perp\mathbf{a}}$ is orthogonal to $\mathbf{a}$, it is also orthogonal to $\mathbf{p} - \mathbf{q}$. Note $\mathbf{b}^{\perp\mathbf{a}} = \mathbf{b} - \mathbf{p}$.
- Hence by the Pythagorean Theorem,

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



Orthogonality helps solve the *fire engine* problem

Fire Engine Lemma:

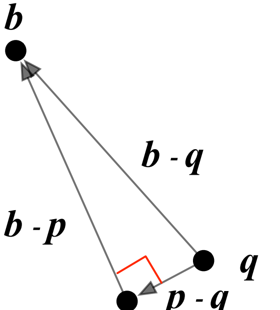
- Let $\mathbf{b}$ be a vector.
- Let $\mathbf{a}$ be a nonzero vector
-

Then $\mathbf{b}^{||\mathbf{a}}$ is the point on the line $\{\alpha\,\mathbf{a}\ :\ \alpha \in \mathbb{R}\}$ that is closest to $\mathbf{b}$, and the distance is $||\mathbf{b}^{\perp\mathbf{a}}||$.

Proof: Let L be the line. Let $\mathbf{p} = \mathbf{b}^{||\mathbf{a}}$. 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 $\mathbf{a}$.
- Hence, since $\mathbf{b}^{\perp\mathbf{a}}$ is orthogonal to $\mathbf{a}$, it is also orthogonal to $\mathbf{p} - \mathbf{q}$. Note $\mathbf{b}^{\perp\mathbf{a}} = \mathbf{b} - \mathbf{p}$.
- Hence by the Pythagorean Theorem,

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



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
- $\mathbf{b}^{\perp\mathbf{a}}$ 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

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

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

$\mathbf{b} = \mathbf{b}^{||\mathbf{a}} + \mathbf{b}^{\perp\mathbf{a}}$ $\qquad$ $\mathbf{b}^{||\mathbf{a}} = \sigma\,\mathbf{a}$ $\qquad$ $\mathbf{b}^{\perp\mathbf{a}}$ 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
```

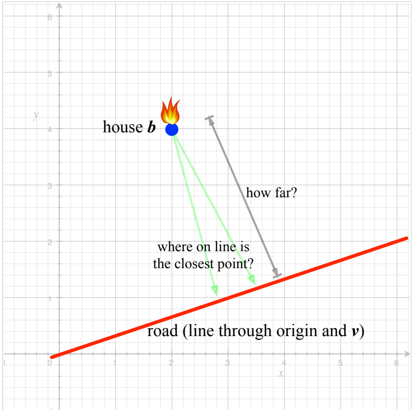
Solution to the *fire engine* problem

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.

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.