Some matrix operations in R

```
A <- matrix(c(4,1,-2,3,-1,0,2,2,-1), nrow=3)     # creates 3-by-3 matrix column-wise
B <- matrix(c(4,1,-2,3,-1,0,2,2,-1), nrow=3, byrow=TRUE)     # creates 3-by-3 matrix row-wise
t(A)    # takes the transpose of matrix A
solve(A)    # numerically computes inverse of A;  A^(-1) DOES NOT WORK!
A %*% B     # computes matrix product of A (left-hand factor) and B (right-hand factor)
A %*% c(3,-1,2)     # computes matrix-vector product; note conversion of vector to 3-by-1

u = c(3,-1,1)
v = c(2,2,8)
dot(u, v)          # takes dot product of u and v
u %*% v            # also takes dot product of u and v
length(u)          # tells how many components are in u
vlength(u)         # computes the magnitude of u, i.e.,  sqrt( dot(u,u) )
```

Projections

Recall, from either MATH 231 or MATH 255, that

- The **span** of a collection **S** of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is the set of all possible sums of rescalings—all **linear combinations**— of those vectors. That is,

$$V \;=\; \text{span}(S) \;:=\; \left\{ c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k \,\middle|\, c_1, \ldots, c_k \in \mathbb{R} \right\}.$$

  In relation to V, the set S is called a **spanning set**. A linear combination

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$$

  can be thought of as a trek through $\mathbb{R}^n$, appending tail-to-head the rescaled vectors $c_j\mathbf{v}_j$ one by one until a destination is reached. V consists of all destinations reachable this way.

  S is hardly the only spanning set for V; you can easily make another from the first:

  - a superset of S: $S \cup \{\mathbf{v}\} \;=\; \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \mathbf{v}\}$, where $\mathbf{v} \in V \setminus S$.
  - a set the size of S: $\{\mathbf{v}_1 - \mathbf{v}_2, \mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_3 \ldots, \mathbf{v}_k\}$,

- A collection $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_\ell\}$ is said to be **linearly independent** if the only way a linear combination

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_\ell\mathbf{u}_\ell$$

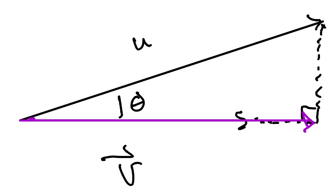
  produces the destination **0** is by taking each $c_1 = c_2 = \cdots = c_\ell = 0$.

- A basis of a vector space V is a linearly independent spanning set for V. There are generally many bases of V, but all consist of the same number of vectors, a number known as the **dimension** of V.

- $\mathbb{R}^n$ is an n-dimensional vector space.

- For any collection $S \subset \mathbb{R}^n$, span(S) is a subspace of $\mathbb{R}^n$, a vector space in its own right living inside $\mathbb{R}^n$.
- There are vector subspaces of all dimensions $d = 0, 1, 2, \ldots, n$ in $\mathbb{R}^n$. The (only) 0-dimensional subspace consists of the single point at the origin, $\{\mathbf{0}\}$. The one-dimensional subspaces consist of rescalings $\{c_1\mathbf{v}_1 \mid c_1 \in \mathbb{R}\}$ of a single nonzero vector $\mathbf{v}_1 \in \mathbb{R}^n$, lines through the origin. The two-dimensional subspaces have the form $\{c_1\mathbf{v}_1 + c_2\mathbf{v}_2 \mid c_1, c_2 \in \mathbb{R}\}$ and consist of linear combinations of two linearly independent vectors $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^n$, planes through the origin. And so it goes, right up until you reach $\mathbb{R}^n$ itself, the only n-dimensional subspace of $\mathbb{R}^n$. This description comprises all subspaces of $\mathbb{R}^n$.
- Vectors have magnitude (length) and direction. We denote the magnitude of $\mathbf{v}$ by $|\mathbf{v}|$, and compute it as the square root
$$|\mathbf{v}| = \sqrt{\mathbf{v}\cdot\mathbf{v}}.$$
The **direction** of a nonzero $\mathbf{v}$ is a vector of length 1 obtained by rescaling $\mathbf{v}$:
$$\frac{1}{|\mathbf{v}|}\mathbf{v}.$$
- Given a subspace W of $\mathbb{R}^n$ and a vector $\mathbf{u} \in \mathbb{R}^n$, there exists a single vector in W closest to $\mathbf{u}$. We call this vector the **projection of u onto** W, or proj$(\mathbf{u} \to W)$. This vector can and will be $\mathbf{u}$ itself if $\mathbf{u}$ comes from W in the first place. The more interesting case, where $\mathbf{u} \in \mathbb{R}^n \setminus \mathbf{W}$, is the focus for the rest of these notes.

When W is a line through the origin, we let $\mathbf{v}$ be a basis vector. Then, proj$(\mathbf{u} \to W)$, or proj$(\mathbf{u} \to \mathbf{v})$, is given by
$$\text{proj}(\mathbf{u} \to \mathbf{v}) = \frac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{v}|^2}\mathbf{v}. \approx \frac{\vec{u}\circ\vec{v}}{|\vec{v}|}\cdot\frac{\vec{v}}{|\vec{v}|}$$



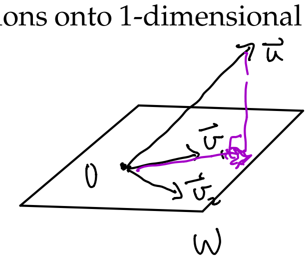
proj$(\vec{u} \to \vec{v})$

Complications arise when the dimension of W reaches two or more.

Example 1: / plane in $\mathbb{R}^3$

Let $W = \text{span}(\{\underbrace{\langle 3,4,0\rangle}_{}^{\vec{v}_1}, \overset{\vec{v}_2}{\langle 2,-2,1\rangle}\})$, and $\mathbf{u} = \langle \underline{3,1,-2}\rangle$.

(a) Find $\underline{\text{proj}(\mathbf{u} \to \langle 3,4,0\rangle)}$ and proj$(\mathbf{u} \to \langle 2,-2,1\rangle)$. These are projections onto 1-dimensional spaces (lines).



$$\text{proj}(\vec{u} \to \vec{v}_1) = \frac{\vec{u}\circ\vec{v}_1}{|\vec{v}_1|^2}\vec{v}_1 = \frac{13}{25}\langle 3,4,0\rangle = \langle \frac{39}{25}, \frac{52}{25}, 0\rangle$$

$$\text{proj}(\vec{u} \to \vec{v}_2) = \langle \frac{4}{9}, \frac{-4}{9}, \frac{2}{9}\rangle$$

(b) Show that $\text{proj}(\mathbf{u} \to W)$ is *not* what one gets by adding $\text{proj}(\mathbf{u} \to \langle 3,4,1\rangle) + \text{proj}(\mathbf{u} \to \langle 2,-2,1\rangle)$.

■

$$\text{proj}(\vec{u} \to W) \overset{?}{=} \text{proj}(\vec{u} \to \vec{v}_1) + \text{proj}(\vec{u} \to \vec{v}_2)$$

$$\vec{z} = \langle 2.004, 1.636, 0.222 \rangle$$

If so, $\vec{u} - \vec{z}$ should be orthogonal to $\vec{z}$

They are not in this case, and usually won't be!

Prescription for finding $\text{proj}(\mathbf{u} \to W)$:

- Find a basis of W: $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k$ We have $(\vec{v}_1, \vec{v}_2)$ already
- Form a matrix $\mathbf{A}$ from the the $\mathbf{w}_j$. Make $\mathbf{w}_1$ the first column of $\mathbf{A}$, $\mathbf{w}_2$ the second column, etc.
- Compute one or both of the following, as desired
  - $\mathbf{x} = (\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{u}$
    The vector $\mathbf{x}$ will have k components, specifying weights $x_1, \ldots, x_k$ such that the linear combination
    $$x_1\mathbf{w}_1 + x_2\mathbf{w}_2 + \cdots + x_k\mathbf{w}_k$$
    is $\text{proj}(\mathbf{u} \to W)$, the closest vector in W to $\mathbf{u}$.
  - $\mathbf{w} = \mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{u}$
    The vector $\mathbf{w}$ is $\text{proj}(\mathbf{u} \to W)$, the closest vector in W to $\mathbf{u}$.

$$A = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 \\ \downarrow & \downarrow \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 4 & -2 \\ 0 & 1 \end{bmatrix}$$

Example 2:

Let $W = \text{span}(\{\langle 3,4,0\rangle, \langle 2,-2,1\rangle\})$, and $\mathbf{u} = \langle 3,1,-2\rangle$. Find $\text{proj}(\mathbf{u} \to W)$.

```
A = matrix( c(3,4,0,2,-2,1), nrow=3 )
u = c(3, 1, -2)
w = A %*% solve( t(A) %*% A ) %*% t(A) %*% u
w
```

```
          [,1]
[1,] 2.3303167
[2,] 1.5022624
[3,] 0.3438914
```

To see that, indeed, $\mathbf{w} = \langle 2.330, 1.502, 1.344\rangle$ is a vector in W:

```
v1 <- c(3,4,0)
v2 <- c(2,-2,1)
x <- solve( t(A) %*% A ) %*% t(A) %*% u     # weights for linear comb
x[1] * v1 + x[2] * v2         # linear combination of v1 and v2 yields w
```

```
[1] 2.3303167 1.5022624 0.3438914
```

To see that $\mathbf{w}$ is as close as possible, we show that $\mathbf{w}$ and $(\mathbf{u} - \mathbf{w})$ are orthogonal. Their dot product is

```
dot(u-w, w)
```

```
[1] 9.992007e-16
```

which, numerically speaking, is zero.

■