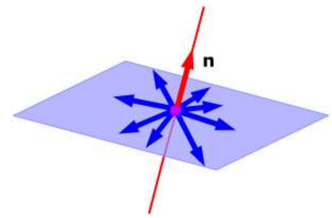
Applied Linear Algebra

Instructor: Jodin Morey   moreyjc@umn.edu

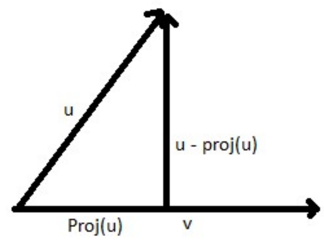
4.4 Orthogonal Projections and Subspaces

Orthogonal Projection

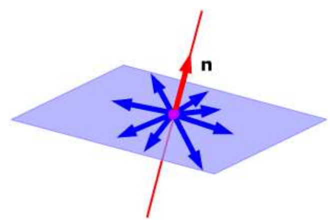
Definition: A vector $\vec{z} \in V$ is said to be *orthogonal to a subspace* $W \subset V$ if it is orthogonal to every vector in W, so $\langle \vec{z}, \vec{w} \rangle = 0$ for all $\vec{w} \in W$.



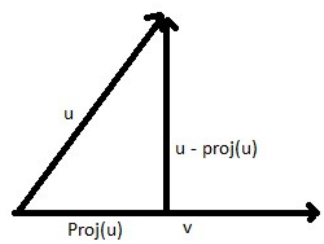
Definition: The *orthogonal projection of* $\vec{u}$ *onto a subspace* W is the element $\vec{w} \in W$ that makes the difference $\vec{z} = \vec{u} - \vec{w} = \vec{u} - Proj_W(u)$ orthogonal to W.
(used in least squares minimization and data fitting)



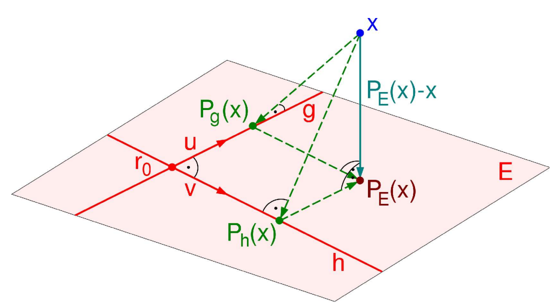
Theorem: Let $\vec{u}_1, \ldots, \vec{u}_n$ be an orthonormal basis for the subspace $W \subset V$. Then the orthogonal projection of $\vec{v} \in V$ onto $\vec{w} \in W$ is given by
$\vec{w} = c_1\vec{u}_1 + \ldots + c_n\vec{u}_n$, where $c_i = \langle \vec{v}, \vec{u}_i \rangle$, $i = 1, \ldots, n$.



Definition: The *orthogonal projection of* $\vec{u}$ *onto a subspace* W is the element $\vec{w} \in W$ that makes the difference $\vec{z} = \vec{u} - \vec{w} = \vec{u} - Proj_W(u)$ orthogonal to W.
(used in least squares minimization and data fitting)



Theorem: Let $\vec{u}_1, \ldots, \vec{u}_n$ be an orthonormal basis for the subspace $W \subset V$. Then the orthogonal projection of $\vec{v} \in V$ onto $\vec{w} \in W$ is given by
$\vec{w} = c_1\vec{u}_1 + \ldots + c_n\vec{u}_n$, where $c_i = \langle \vec{v}, \vec{u}_i \rangle$, $i = 1, \ldots, n$. $(*)$



Proof: Since $\vec{u}_1, \ldots, \vec{u}_n$ form a basis of the subspace, the orthogonal projection element $\vec{w}$ must be some linear combination thereof: $\vec{w} = c_1\vec{u}_1 + \ldots + c_n\vec{u}_n$.

Definition above requires the difference $\vec{z} = \vec{v} - \vec{w}$ be orthogonal to W, and, as noted above, it suffices to check orthogonality to the basis vectors. By our orthonormality assumption:

$$0 = \langle \vec{z}, \vec{u}_i \rangle = \langle \vec{v} - \vec{w}, \vec{u}_i \rangle = \langle \vec{v} - c_1\vec{u}_1 - \ldots - c_n\vec{u}_n, \vec{u}_i \rangle$$

$$= \langle \vec{v}, \vec{u}_i \rangle - c_1\langle \vec{u}_1, \vec{u}_i \rangle - \ldots - c_n\langle \vec{u}_n, \vec{u}_i \rangle = \langle \vec{v}, \vec{u}_i \rangle - c_i.$$

The coefficients $c_i = \langle \vec{v}, \vec{u}_i \rangle$ of the orthogonal projection $\vec{w}$ are therefore uniquely prescribed by the orthogonality requirement, which thereby proves its uniqueness. ■

More generally, with an orthogonal basis, the previous argument will show the orthogonal projection of $\vec{v}$ onto W is given by: $\vec{w} = a_1\vec{v}_1 + \ldots + a_n\vec{v}_n$, where $a_i = \frac{\langle \vec{v}, \vec{v}_i \rangle}{|\vec{v}_i|^2}$, $i = 1, \ldots, n$.

Concretely: Consider the plane $W \subset \mathbb{R}^3$ spanned by orthogonal vectors $\vec{v}_1 = (1,-2,1)$ and $\vec{v}_2 = (1,1,1)$.

According to the formula above, the orthogonal projection of $\vec{v} = (1,0,0)$ onto W is

$$\vec{w} = \frac{\langle \vec{v}, \vec{v}_1 \rangle}{|\vec{v}_1|^2}\vec{v}_1 + \frac{\langle \vec{v}, \vec{v}_2 \rangle}{|\vec{v}_2|^2}\vec{v}_2 = \tfrac{1}{6}(1,-2,1) + \tfrac{1}{3}(1,1,1) = \left(\tfrac{1}{2}, 0, \tfrac{1}{2}\right).$$

Alternatively, we can replace $\vec{v}_1, \vec{v}_2$ by the orthonormal basis: $\vec{u}_1 = \frac{\vec{v}_1}{|\vec{v}_1|} = \left(\frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)$ and

$$\vec{u}_2 = \frac{\vec{v}_2}{|\vec{v}_2|} = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right).$$

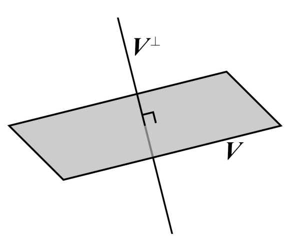
Then, using the orthonormal version $(*)$, $\vec{w} = \langle \vec{v}, \vec{u}_1 \rangle \vec{u}_1 + \langle \vec{v}, \vec{u}_2 \rangle \vec{u}_2$

$$= \frac{1}{\sqrt{6}}\left(\frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right) + \frac{1}{\sqrt{3}}\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) = \left(\tfrac{1}{2}, 0, \tfrac{1}{2}\right).$$

❗ The orthogonal projection formula includes the orthogonal basis formula as a special case.

Orthogonal Subspaces

Subspaces can be orthogonal.



$V \perp V^\perp$

Definition: Two subspaces $W, Z \subset V$ are called *orthogonal subspaces* if every vector in W is orthogonal to every vector in Z.

Lemma: If $\vec{w}_1, \ldots, \vec{w}_k$ span W, and $\vec{z}_1, \ldots, \vec{z}_l$ span Z, then W and Z are orthogonal subspaces **iff** $\langle \vec{w}_i, \vec{z}_j \rangle = 0$ for all $i = 1, \ldots, k$ and $j = 1, \ldots, l$.

So $W \perp Z$ **iff** every basis element of W is orthogonal to every basis element of Z.

Example: Let W be the span of $\vec{w}_1 = (1, -2, 0, 1)^T$ and $\vec{w}_2 = (3, -5, 2, 1)^T$, and let Z be the span of $\vec{z}_1 = (3, 2, 0, 1)^T$ and $\vec{z}_2 = (1, 0, -1, -1)^T$. Are W and Z are orthogonal?

You can check that $\vec{w}_1 \cdot \vec{z}_1 = \vec{w}_1 \cdot \vec{z}_2 = \vec{w}_2 \cdot \vec{z}_1 = \vec{w}_2 \cdot \vec{z}_2 = 0$.
Therefore, W and Z are orthogonal 2D subspaces of $\mathbb{R}^4$ with respect to Euclidean dot product.

Definition: The *orthogonal complement of a subspace* $W \subset V$, denoted $W^\perp$, is defined as the set of all vectors that are orthogonal to W: so $W^\perp = \{\vec{v} \in V \mid \langle \vec{v}, \vec{w} \rangle = 0 \text{ for all } \vec{w} \in W\}$. (depends upon choice of inner product)

Concretely: Let $W = \{(t, 2t, 3t)^T : t \in \mathbb{R}\}$ be the line (1D subspace) in the direction of $\vec{w}_1 = (1, 2, 3)^T \in \mathbb{R}^3$.

Under dot product, its orthogonal complement $W^\perp = \vec{w}_1^\perp$ is the plane passing through the origin having normal vector $\vec{w}_1$.

In other words, $\vec{z} = (x, y, z)^T \in W^\perp$ **iff** $\vec{z} \cdot \vec{w}_1 = x + 2y + 3z = 0$. Thus, $W^\perp$ is characterized as the solution space

of the homogeneous linear equation above, or equivalently, the kernel of $\mathbf{A} = \vec{w}_1^T = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$.

We can write the general solution in the form: $\vec{z} =$??

$$\begin{bmatrix} -2y - 3z \\ y \\ z \end{bmatrix} = y \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} = y\vec{z}_1 + z\vec{z}_2,$$

where y, z are the free variables. So $\vec{z}_1 = (-2, 1, 0)^T$ and $\vec{z}_2 = (-3, 0, 1)^T$ form a (non-orthogonal) basis for $W^\perp$.

Proposition: Suppose that $W \subset V$ is a finite dimensional subspace of an inner product space.
Then every vector $\vec{v} \in V$ can be *uniquely* decomposed into $\vec{v} = \vec{w} + \vec{z}$, where $\vec{w} \in W$ and $\vec{z} \in W^\perp$.

Proof: Let $\vec{w} \in W$ be orthogonal projection of $\vec{v}$ onto W. Then, $\vec{z} = \vec{v} - \vec{w}$ is, by definition, orthogonal to W and hence belongs to $W^\perp$.

Note that $\vec{z}$ can be viewed as orthogonal projection of $\vec{v}$ onto complementary subspace $W^\perp$ (if it's finite dimensional).

Now we just need uniqueness. If we are given two such decompositions, $\vec{v} = \vec{w} + \vec{z} = \vec{w}' + \vec{z}'$, then $\vec{w} - \vec{w}' = \vec{z}' - \vec{z}$.

The left-hand side of this equation lies in W, while the right-hand side belongs to $W^{\perp}$.

But, as we already noted, the only vector that belongs to both W and $W^{\perp}$ is the zero vector.

Thus, $\vec{w} - \vec{w}' = \vec{0} = \vec{z}' - \vec{z}$, so $\vec{w} = \vec{w}'$ and $\vec{z} = \vec{z}'$, which proves uniqueness. ■

Proposition: If $W \subset V$ is a subspace with $\dim W = n$ and $\dim V = m$, then $\dim W^{\perp} = m - n$.

Proposition: If W is a finite dimensional subspace of an inner product space, then $(W^{\perp})^{\perp} = W$.

Orthogonality of the Fundamental Matrix Subspaces and the Fredholm Alternative

Theorem: Let $\mathbf{A}$ be a real $m \times n$ matrix. Then $\mathbf{A}$'s kernel and coimage are orthogonal complements as subspaces of $\mathbb{R}^n$ under the dot product. Also, $\mathbf{A}$'s cokernel and image are orthogonal complements in $\mathbb{R}^m$:
$\ker \mathbf{A} = (coimg\,\mathbf{A})^{\perp} \subset \mathbb{R}^n, \qquad co\ker \mathbf{A} = (img\,\mathbf{A})^{\perp} \subset \mathbb{R}^m.$

Proof: $\vec{x} \in \mathbb{R}^n$ lies in $\ker \mathbf{A}$ **iff** $\mathbf{A}\vec{x} = \vec{0}$. Observe, the i^{th} entry of $\mathbf{A}\vec{x}$ equals product of i^{th} row $\vec{r}_i^{\,T}$ of $\mathbf{A}$ and $\vec{x}$.

But $\vec{r}_i^{\,T}\vec{x} = \vec{r}_i \cdot \vec{x} = 0$, **iff** $\vec{x}$ is orthogonal to $\vec{r}_i$. Therefore, $\vec{x} \in \ker \mathbf{A}$ **iff** $\vec{x}$ is orthogonal to all rows of $\mathbf{A}$.

Since these rows span $coimg\,\mathbf{A}$, this is equivalent to $\vec{x}$ lying in $(coimg\,\mathbf{A})^{\perp}$, which proves the first statement.

Orthogonality of the image and cokernel follows by the same argument applied to $\mathbf{A}^T$. ■

Therefore (given the rank/nullity thm):
Theorem (*Fredholm Alternative, "FH"*): $\mathbf{A}\vec{x} = \vec{b}$ has a solution **iff** $\vec{b}$ is orthogonal to the cokernel of $\mathbf{A}$.

Indeed, since $\mathbf{A}$'s cokernel and image are orthogonal compliments, and $\dim(coker\mathbf{A}) + \dim(im\mathbf{A}) = m$,

$\vec{b} \neq 0$ must either be in the cokernel or the image. If $\vec{b}$ is in the image, by the previous theorem it is orthogonal to the cokernel.

If it is in the cokernel, it is orthogonal to the image, and the system has no solution.

Thus, the compatibility conditions for $\mathbf{A}\vec{x} = \vec{b}$ are: $\vec{y} \cdot \vec{b} = 0$ for every $\vec{y}$ such that $\mathbf{A}^T\vec{y} = \vec{0}$.

Or more efficiently, one can check that $\vec{b}$ is orthogonal with respect to the cokernel's basis vectors.

Concretely: Let's determine compatibility conditions for:

$$x_1 - x_2 + 3x_3 = b_1, \qquad -x_1 + 2x_2 - 4x_3 = b_2, \qquad 2x_1 + 3x_2 + x_3 = b_3, \qquad x_1 + 2x_3 = b_4.$$

By FH, need to solve: $\mathbf{A}^T\vec{y} = \vec{0}$ where $\mathbf{A} = \begin{bmatrix} 1 & -1 & 3 \\ -1 & 2 & -4 \\ 2 & 3 & 1 \\ 1 & 0 & 2 \end{bmatrix}$.

$$\mathbf{A}^T = \begin{bmatrix} 1 & -1 & 2 & 1 \\ -1 & 2 & 3 & 0 \\ 3 & -4 & 1 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 7 & 2 \\ 0 & 1 & 5 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

y_4, y_3 are free. $y_2 = -5y_3 - y_4$ and $y_1 = -7y_3 - 2y_4$ or

$$\vec{y} = \left(-7y_3 - 2y_4, \; -5y_3 - y_4, \; y_3, \; y_4\right) = y_3(-7,-5,1,0) + y_4(-2,-1,0,1).$$

Therefore, by the Fredholm alternative we need to find vectors which are orthogonal to this space.

Dotting $\vec{b}$ with each basis vectors gives compatibility conditions: $-7b_1 - 5b_2 + b_3 = 0$ and $-2b_1 - b_2 + b_4 = 0$.

If you were to do the Gaussian elimination on $\left[\mathbf{A} \,|\, \vec{b}\right]$, you would find the same compatibility conditions.

❗ **Recharacterizing A**: Let $\vec{b} = \mathbf{A}\vec{x} \in \mathbb{R}^m$. Since kernel and coimage of $\mathbf{A}$ are orthogonal complements in domain $\mathbb{R}^n$, we can *uniquely* decompose $\vec{x} = \vec{w} + \vec{z}$, where $\vec{w} \in coimg\mathbf{A}$, while $\vec{z} \in \ker\mathbf{A}$.

And since $\mathbf{A}\vec{z} = \vec{0}$, it must be that $\vec{b} = \mathbf{A}\vec{x} = \mathbf{A}(\vec{w} + \vec{z}) = \mathbf{A}\vec{w}$.

Therefore, multiplying an $\vec{x}$ in the domain by $\mathbf{A}$ can be seen as 1st projecting $\vec{x}$ onto the coimage ($\vec{x} \to \vec{w}$), and then mapping $\vec{w}$ to the $img\mathbf{A}$: $\mathbf{A}\vec{x} = \mathbf{A}\vec{w} = \vec{b}$. This provides a 1-1 correspondence between $img\mathbf{A}$ & $coimg\mathbf{A}$.

Also, observe that if $\mathbf{A}$ has rank r, then both $img\mathbf{A}$ & $coimg\mathbf{A}$ are r-dimensional, albeit of different vector spaces.

Therefore:

Theorem: Multiplication by $\mathbf{A}^{m\times n}$ of rank r defines a one-to-one correspondence between the r-dimensional subspaces $img\mathbf{A} \subset \mathbb{R}^m$ and $coimg\mathbf{A} \subset \mathbb{R}^n$. Moreover, if $\vec{v}_1, \ldots, \vec{v}_r$ forms a basis of $coimg\,\mathbf{A}$, then their images under $\mathbf{A}$: $\mathbf{A}\vec{v}_1, \ldots, \mathbf{A}\vec{v}_r$ form a basis for $img\,\mathbf{A}$.

Theorem: A compatible $\mathbf{A}\vec{x} = \vec{b}$ with $\vec{b} \in img\,\mathbf{A} = (co\ker\mathbf{A})^{\perp}$ has a *unique* solution $\vec{w} \in coimg\,\mathbf{A}$ satisfying $\mathbf{A}\vec{w} = \vec{b}$. This general solution is $\vec{x} = \vec{w} + \vec{z}$, where $\vec{z} \in \ker\mathbf{A}$. The particular solution $\vec{w} \in coimg\,\mathbf{A}$ has the smallest Euclidean norm of all possible solutions: $|\vec{w}| \le |\vec{x}|$ whenever $\mathbf{A}\vec{x} = \vec{b}$.

Partial Proof: Does $\vec{b} \in img\mathbf{A}$ corresponds to a unique $\vec{w} \in coimg\mathbf{A}$?

Indeed, if $\vec{w}, \tilde{\vec{w}} \in coimg\mathbf{A}$ satisfy $\vec{b} = \mathbf{A}\vec{w} = \mathbf{A}\tilde{\vec{w}}$, then $\mathbf{A}(\vec{w} - \tilde{\vec{w}}) = \vec{0}$, and hence $\vec{w} - \tilde{\vec{w}} \in \ker\mathbf{A}$.

But, since kernel & coimage are orthogonal complements, the only vector that belongs to both is $\vec{0}$, and hence $\vec{w} = \tilde{\vec{w}}$.

And, since the coimage & kernel are orthogonal subspaces, the norm of a general solution ($\vec{x} = \vec{w} + \vec{z}$) is:

$$|\vec{x}|^2 = |\vec{w} + \vec{z}|^2 = |\vec{w}|^2 + 2\vec{w} \cdot \vec{z} + |\vec{z}|^2 = |\vec{w}|^2 + |\vec{z}|^2 \ge |\vec{w}|^2, \text{ with equality } \textbf{iff } \vec{z} = \vec{0}. \qquad \blacksquare$$

Example: Find the orthogonal complement $W^{\perp}$ of the subspace $W \subset \mathbb{R}^3$ spanned by

$$\vec{v}_1 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}.$$ What is the dimension of $W^{\perp}$?

Observe that $2\vec{v}_3 - \vec{v}_2 = \vec{v}_1$. However, by observation $\vec{v}_3$ is linearly independent from $\vec{v}_2$: W is 2D.

Therefore, since we are in $\mathbb{R}^3$, the orthogonal complement $W^{\perp}$ must be 1D.

In particular, we need $\vec{w} = \langle w_1, w_2, w_3 \rangle$ such that $\langle \vec{w}, \vec{v}_2 \rangle = \langle \vec{w}, \vec{v}_3 \rangle = 0$.

Or: $-2w_1 + 3w_2 + w_3 = 0$ and $-w_1 + 2w_2 = 0$.

From the second equation: $w_1 = 2w_2$, and from the first equation:

$$-2(2w_2) + 3w_2 + w_3 = w_3 - w_2 = 0 \text{ or } w_2 = w_3.$$

So the orthogonal complement is spanned by $\vec{w} = (2, 1, 1)$. Checking our work, one sees that $\langle \vec{w}, \vec{v}_2 \rangle = \langle \vec{w}, \vec{v}_3 \rangle = 0$.

For the homework, given the bases for two subspaces, how can we determine if these subspaces are orthogonal?