Orthogonal projection of vector $\vec{v}$ on direction of vector $\vec{u}$

If $\vec{u}$ and $\vec{v}$ are vectors in $\mathbb{R}^n$ and $\vec{u} \neq \vec{0}$,
there exist unique vectors $\vec{v}\perp$ and $\vec{v}\|$ satisfying the following:

- $\vec{v} = \vec{v}\perp + \vec{v}\|$
- $\vec{u} \cdot \vec{v}\perp = 0$
- $\vec{v}\| = s\,\vec{u}$, where s is a scalar

Then

$$\vec{u} \cdot \vec{v} =$$
$$\vec{u} \cdot (\vec{v}\perp + \vec{v}\|) =$$
$$\vec{u} \cdot \vec{v}\perp + \vec{u} \cdot (s\,\vec{u}) =$$
$$0 + s\,(\vec{u} \cdot \vec{u})$$

$$\text{so } s = \frac{\vec{u} \cdot \vec{v}}{\vec{u} \cdot \vec{u}}$$

- $\vec{v}\| = \left(\dfrac{\vec{u} \cdot \vec{v}}{\vec{u} \cdot \vec{u}}\right) \vec{u}$
- $\vec{v}\perp = \vec{v} - \left(\dfrac{\vec{u} \cdot \vec{v}}{\vec{u} \cdot \vec{u}}\right) \vec{u}$

Orthogonal projection of vector $\vec{v}$ on columns of $U = [\ \vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r\ ]$

This is the same concept as the projection of $\vec{v}$ on one vector $\vec{u}$
The only difference is that we replace one direction $\vec{u}$ by r directions $\vec{u}_1, \ldots, \vec{u}_r$

Suppose
- $U = [\, \vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r \,]$
- $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r \in \mathbb{R}^n$ and are linearly independent
- $\vec{v} \in \mathbb{R}^n$

There exist unique vectors $\vec{v}\perp$ and $\vec{v}\|$ such that:
- $\vec{v} = \vec{v}\perp + \vec{v}\|$
- $U^T \vec{v}\perp = \vec{0}$

(in other words, $\vec{v}\perp$ is orthogonal to every column of U or every row of U^T)
- $\vec{v}\| = U \vec{c}$, where $\vec{c} \in \mathbb{R}^r$

($\vec{c}$ is similar to scalar 's' used on previous page)

Then
$$U^T \vec{v} = U^T (\vec{v}\perp + \vec{v}\|) = U^T \vec{v}\perp + U^T (U \vec{c}) = \vec{0} + (U^T U) \vec{c}$$

So $\vec{c}$ satisfies
$$(U^T U) \vec{c} = U^T \vec{v}$$

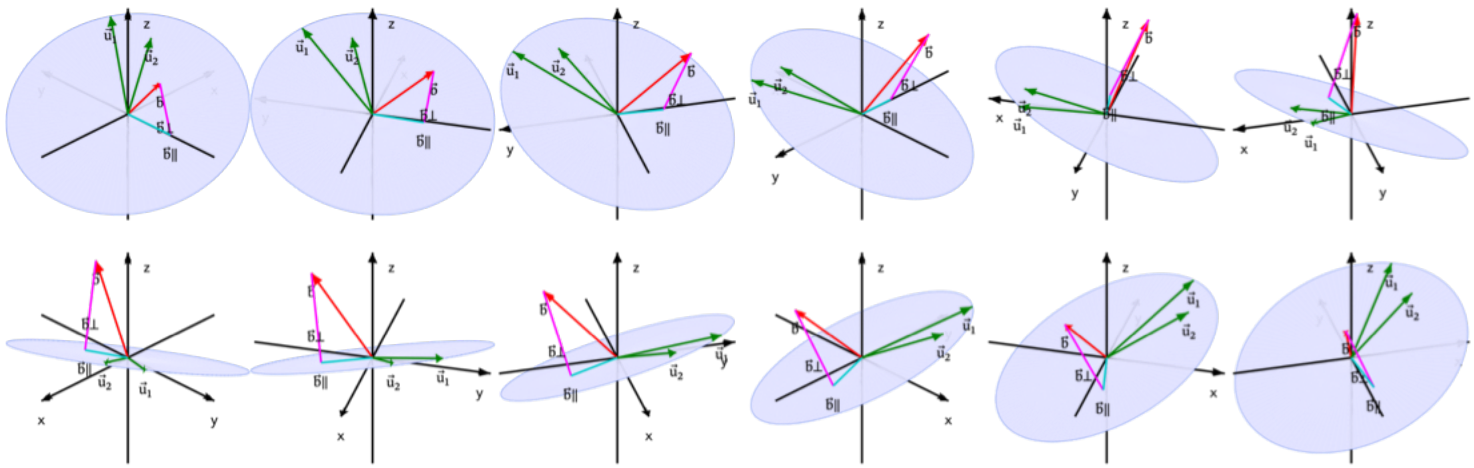
Therefore
$$\vec{c} = (U^T U)^{-1} U^T \vec{v}$$

And the projection of $\vec{v}$ on columns of U is
$$\vec{v}\| = U \vec{c} = U (U^T U)^{-1} U^T \vec{v}$$

So the projection matrix onto span($\vec{u}_1, \ldots, \vec{u}_r$) is
$$P = U (U^T U)^{-1} U^T$$

The same orthogonality principle appeared earlier in the 'Four subspaces' chapter
where we were solving an inconsistent system $A \vec{x} = \vec{b}$
where $\vec{b}$ was outside the column space(A)

and used the same fundamental conditions and derivation sequence:



The same structure in three equivalent forms:

Projection, vector form	Projection, matrix form	From 'Four subspaces'
$\vec{v} = \vec{v}\perp + \vec{v}\Vert$	$\vec{v} = \vec{v}\perp + \vec{v}\Vert$	$\vec{b} = \vec{b}\perp + \vec{b}\Vert$
$\vec{u} \cdot \vec{v}\perp = 0$	$U^T \vec{v}\perp = \vec{0}$	$A^T \vec{b}\perp = \vec{0}$
$\vec{v}\Vert = s\,\vec{u}$	$\vec{v}\Vert = U\,\vec{c}$	$\vec{b}\Vert = A\,\vec{x}$
$\vec{u} \cdot \vec{v} =$ $\vec{u} \cdot (\vec{v}\perp + \vec{v}\Vert) =$ $\vec{u} \cdot \vec{v}\Vert =$ $s\,\vec{u} \cdot \vec{u}$	$U^T \vec{v} =$ $U^T (\vec{v}\perp + \vec{v}\Vert) =$ $U^T \vec{v}\Vert =$ $U^T U\,\vec{c}$	$A^T \vec{b} =$ $A^T (\vec{b}\perp + \vec{b}\Vert) =$ $A^T \vec{b}\Vert =$ $A^T A\,\vec{x}$
$s = \dfrac{\vec{u} \cdot \vec{v}}{\vec{u} \cdot \vec{u}}$	$\vec{c} = (U^T U)^{-1} U^T \vec{v}$	$\vec{x} = (A^T A)^{-1} A^T \vec{b}$

While the 'Four subspaces' page stopped at solving the normal equation

$$\vec{x} = (A^T A)^{-1} A^T \vec{b}$$

this one takes one more step to derive the corresponding projection matrix

$$P = U\,(U^T U)^{-1}\,U^T$$

that inputs $\vec{v}$ and outputs U $\vec{c}$:

$$P \vec{v} = U (U^T U)^{-1} U^T \vec{v} = U \vec{c}$$

Orthogonal vs general projection (optional reading)

To define projection, we need 2 complementary subspaces:

① Subspace to project onto, denote as $\mathcal{P}$

② Subspace that contains the residual, denote as $\mathcal{R}$

Recall definition of complementary:

- $\mathcal{P} \cap \mathcal{R} = \vec{0}$ (condition 1)
- $\mathcal{P} \oplus \mathcal{R} = \mathbb{R}^m$ (condition 2)

Suppose we are projecting onto columns of a full-column rank m×n matrix U:

$$\mathcal{P} = \text{col}(U)$$

We know that col(U) has an orthogonal complement:

$$\ell\text{-null}(U) = \text{null}(U^T),$$

and that defines $\mathcal{R}$ for orthogonal projection

Next, we define $\mathcal{R}$ for general projection:

Suppose there is another full-column rank m×n matrix S

Assume $S^T U$ is invertible

To prove condition 1, use the rank formula for matrix product:

$$\text{rank}(S^T U) = \text{rank}(U) - \dim\left(\text{col}(U) \cap \text{null}(S^T)\right)$$

Since U has full column rank, rank(U) = n
Since $S^T U$ is invertible, $\text{rank}(S^T U) = n$

↓

$$n = n - \dim\left(\text{col}(U) \cap \text{null}(S^T)\right)$$

↓

$$\dim\left(\text{col}(U) \cap \text{null}(S^T)\right) = 0$$

↓

$\text{col}(U) \cap \text{null}(S^T) = \vec{0}$
(condition 1)

To prove condition 2:

$$\dim\left(\text{col}(U)\right) = n$$

because U has n linearly independent columns

Also $\text{rank}(S^T) = \text{rank}(S) = n$
so by rank-nullity,

$$\dim\left(\text{null}(S^T)\right) = m - n$$

Therefore

$$\dim\left(\text{col}(U)\right) + \dim\left(\text{null}(S^T)\right) = n + (m - n) = m$$

Together with condition 1,
$\text{col}(U) \oplus \text{null}(S^T) = \mathbb{R}^m$
(condition 2)

So $\mathscr{R} = \text{null}(S^T) = \ell\text{-null}(S)$ defines a valid complementary subspace

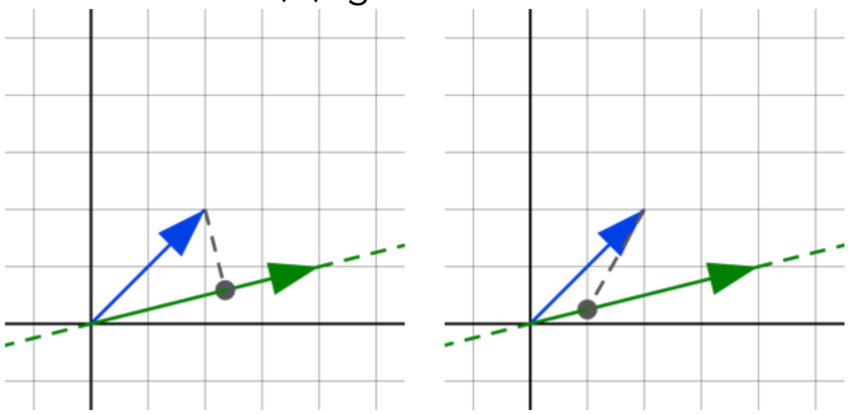
Orthogonal projection is the special case S = U

Comparison table,
including side-by-side derivation of projection matrix

	orthogonal	general
subspace $\mathcal{P}$	col(U)	col(U)
subspace $\mathcal{R}$	null(U^T)	null(S^T) where $S^T U$ is invertible
space decomposition	col(U) ⊕ null(U^T) same as col(U) ⊕ ℓ-null(U)	col(U) ⊕ null(S^T) same as col(U) ⊕ ℓ-null(S)
Left-multiply to eliminate residual	Left-multiply by U^T so that $U^T \vec{v}\perp = \vec{0}$	Left-multiply by S^T so that $S^T \vec{r} = \vec{0}$
Decomposition of $\vec{v}$	$\vec{v} = (\vec{v}\| \in$ col(U)) + ($\vec{v}\perp \in$ ℓ-null(U))	$\vec{v} = (\vec{v}\| \in$ col(U)) + ($\vec{r} \in$ ℓ-null(S))
$\vec{v}\|$	$\vec{v}\| = U\,\vec{c}$	$\vec{v}\| = U\,\vec{d}$
Residual direction	$\vec{v}\perp \in$ ℓ-null(U) ↓ $U^T \vec{v}\perp = \vec{0}$	$\vec{r} \in$ ℓ-null(S) ↓ $S^T \vec{r} = \vec{0}$
Left-multiplication eliminates residual	$U^T \vec{v} = U^T (\vec{v}\perp + \vec{v}\|)$ $= U^T \vec{v}\perp + U^T \vec{v}\|$ $= \vec{0} + U^T U\,\vec{c}$ $= U^T U\,\vec{c}$	$S^T \vec{v} = S^T (\vec{r} + \vec{v}\|)$ $= S^T \vec{r} + S^T \vec{v}\|$ $= \vec{0} + S^T U\,\vec{d}$ $= S^T U\,\vec{d}$
Solve for coefficients	$\vec{c} = (U^T U)^{-1} U^T \vec{v}$	$\vec{d} = (S^T U)^{-1} S^T \vec{v}$
Projection formula	$\vec{v}\| = U\,\vec{c}$ $= U (U^T U)^{-1} U^T \vec{v}$	$\vec{v}\| = U\,\vec{d}$ $= U (S^T U)^{-1} S^T \vec{v}$
Projection matrix	$P = U (U^T U)^{-1} U^T$	$P = U (S^T U)^{-1} S^T$
Residual condition	residual ⊥ col(U)	residual ∈ ℓ-null(S)
Uniqueness	unique $\vec{v}\|$	depends on choice of S
Utility	Best solution to $A\,\vec{x} = \vec{b}$ (minimizes $\|A\,\vec{x} - \vec{b}\|$)	Computer graphics

Visual examples of orthogonal (l) & general (r) projections in $\mathbb{R}^2$

- $\vec{v}$: blue vector
- $\vec{u}_1$: green vector
- $U = [\ \vec{u}_1\ ]$
- col(U): green dotted line



Idempotency of P

P is idempotent which means
it acts once even when applied repeatedly:

① P (orthogonal projection) = $U\,(U^T U)^{-1}\,U^T$

$$P^2 = U\,(U^T\,U)^{-1}\,U^T\,U\,(U^T\,U)^{-1}\,U^T =$$

$$U\,(U^T\,U)^{-1}\,\left((U\,U)\,(U^T\,U)^{-1}\right)\,U^T =$$

$$U (U^T U)^{-1} U^T$$

② P (general projection) = $U (S^T U)^{-1} S^T$

$$P^2 = U (S^T U)^{-1} S^T U (S^T U)^{-1} S^T =$$

$$U (S^T U)^{-1} \left((S^T U) (S^T U)^{-1}\right) S^T =$$

$$U (S^T U)^{-1} S^T$$