1. Analytic Geometry

1.1 Inner Product

1.1.1 Dot Product

Let be $u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ $and \;\; v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ in $\mathbb{R}^n$. The scalar product or the dot product is given by:

$$u \cdot v = u^T v = [u_1 \quad u_2 \quad \cdots \quad u_n] \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n.$$

1.1.2 General Inner Products

Definition: Let be V a vector space, a mapping $\Omega : \; V \; \times V \; \to \mathbb{R}$ such that:

- $\Omega(\alpha x + \beta y, z) = \alpha\Omega(x, z) + \beta\Omega(\alpha y, z)$, $\alpha, \beta \in \mathbb{R}$ and $x, y, z \in V$;
- $\Omega(x, \alpha y + \beta z) = \alpha\Omega(x, y) + \beta\Omega(x, z)$, $\alpha, \beta \in \mathbb{R}$ and $x, y, z \in V$ (bilinear mapping);
- $\forall x \in V\backslash\{0\}: \Omega(x, x) > 0, \;\; \Omega(0,0) = 0$ (positive definite);
- $\Omega(x, y) = \Omega(y, x), \forall x, y \in V$ (symmetric).

It is called an *inner product* on V . We typically write instead $\Omega(x, y)$.

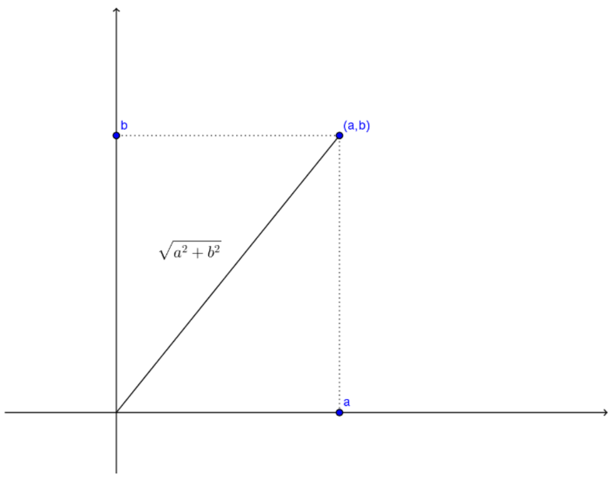
The pair $(V, <\cdot,\cdot>)$ is called an inner product space or (real) vector space inner product space

vector space with *inner product*. If we use the dot product then the space $(V, <\cdot,\cdot>)$ is called an *Euclidean vector space*.

1.1.3 Lengths and Distances

Lengths

Let $u = \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2$, by Pythagoras the length of u is $\sqrt{a^2 + b^2} = \sqrt{u^T u}$



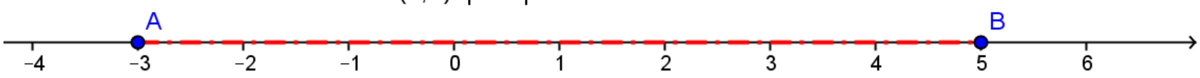
We can define a norm through the inner product

$$\|u\| = \sqrt{u^T u}$$

However, not every norm is induced by an inner product.

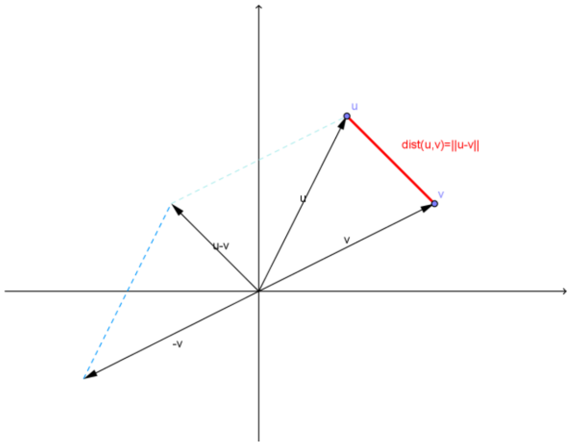
Distance

- $\mathbb{R}$:



$$Dist(a, b) = |a - b|$$

$\mathbb{R}^2$:



Definition 3.6 (Distance and Metric)**.** Consider an inner product space $(V, \langle\cdot,\cdot\rangle)$. Then

$$d(\boldsymbol{x}, \boldsymbol{y}) := \|\boldsymbol{x} - \boldsymbol{y}\| = \sqrt{\langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y}\rangle} \tag{3.21}$$

is called the *distance* between $\boldsymbol{x}$ and $\boldsymbol{y}$ for $\boldsymbol{x}, \boldsymbol{y} \in V$. If we use the dot product as the inner product, then the distance is called *Euclidean distance*.

The mapping

$$d : V \times V \to \mathbb{R} \tag{3.22}$$
$$(\boldsymbol{x}, \boldsymbol{y}) \mapsto d(\boldsymbol{x}, \boldsymbol{y}) \tag{3.23}$$

is called a *metric*.

A metric d satisfies the following:

1. d is *positive definite*, i.e., $d(\boldsymbol{x}, \boldsymbol{y}) \geqslant 0$ for all $\boldsymbol{x}, \boldsymbol{y} \in V$ and $d(\boldsymbol{x}, \boldsymbol{y}) = 0 \iff \boldsymbol{x} = \boldsymbol{y}$.
2. d is *symmetric*, i.e., $d(\boldsymbol{x}, \boldsymbol{y}) = d(\boldsymbol{y}, \boldsymbol{x})$ for all $\boldsymbol{x}, \boldsymbol{y} \in V$.
3. *Triangle inequality:* $d(\boldsymbol{x}, \boldsymbol{z}) \leqslant d(\boldsymbol{x}, \boldsymbol{y}) + d(\boldsymbol{y}, \boldsymbol{z})$ for all $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$.

1.2.Orthogonality

1.2.1 Angles

Definition Let be u, v vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$ then $\cos\alpha = \frac{<u,v>}{\|u\|\|v\|} = \frac{u^T v}{\|u\|\|v\|}$.

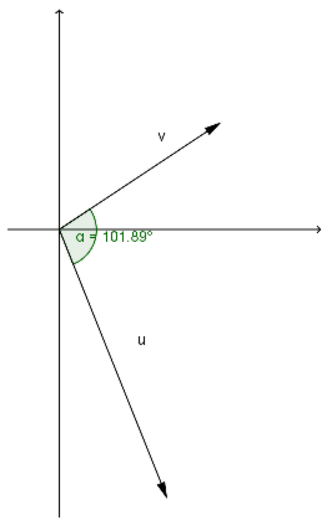
Let us compute the angle between, $u = \begin{bmatrix} 2 \\ -5 \end{bmatrix}$ and $v = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$.

$$\cos\alpha = \frac{2 \times 3 - 5 \times 2}{\sqrt{2^2 + (-5)^2}\sqrt{3^2 + 2^2}}$$

$$\cos\alpha = -\frac{4}{19{,}42}$$

Hence

$$\alpha \approx 101{,}8864766 \ldots$$

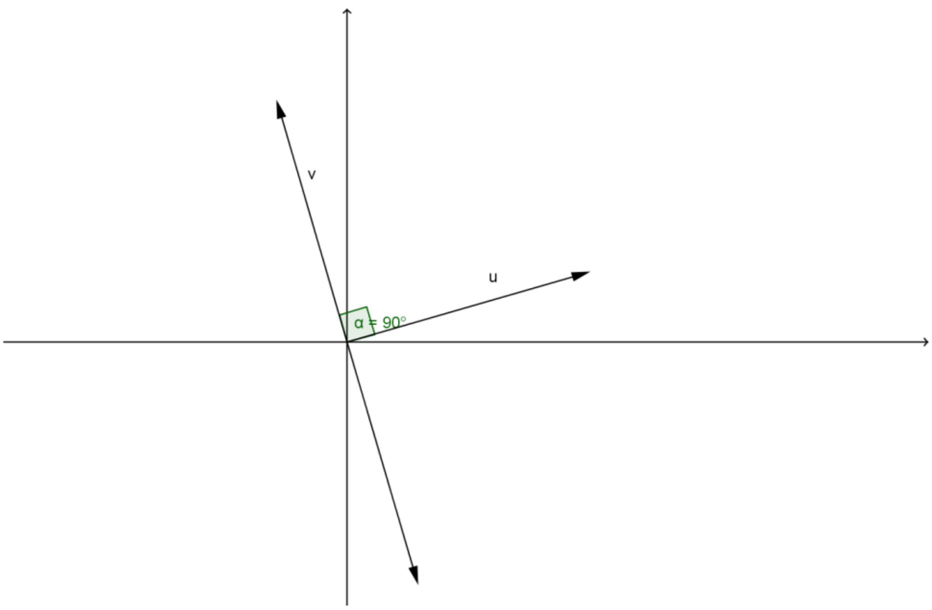


1.2.2 Orthogonal vectors

Definition 3.7 (Orthogonality). Two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ are *orthogonal* if and only if $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$, and we write $\boldsymbol{x} \perp \boldsymbol{y}$. If additionally $\|\boldsymbol{x}\| = 1 = \|\boldsymbol{y}\|$, i.e., the vectors are unit vectors, then $\boldsymbol{x}$ and $\boldsymbol{y}$ are *orthonormal*.

An implication of this definition is that the $\mathbf{0}$-vector is orthogonal to every vector in the vector space.

Remark. Orthogonality is the generalization of the concept of perpendicularity to bilinear forms that do not have to be the dot product. In our context, geometrically, we can think of orthogonal vectors as having a right angle with respect to a specific inner product. ◇



$dist(u, -v) = dist(u, v)$, but

- $dist(u, -v)^2 = \|u - (-v)\|^2 = <u + v, u + v> = <u, u> + <u, v> + <v, u> + <v, v> = \|u\|^2 + \|v\|^2 + 2(u \cdot v)$
- $dist(u, v)^2 = \|u - v\|^2 = <u - v, u - v> = <u, u> - <u, v> - <v, u> + <v, v> = \|u\|^2 + \|v\|^2 - 2(u \cdot v)$

Then $dist(u, -v) = dist(u, v)$,if and only if se $<u, v> = 0$.

Definition 3.8 (Orthogonal Matrix). A square matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is an *orthogonal matrix* if and only if its columns are orthonormal so that

$$\boldsymbol{A}\boldsymbol{A}^\top = \boldsymbol{I} = \boldsymbol{A}^\top\boldsymbol{A}\,, \tag{3.29}$$

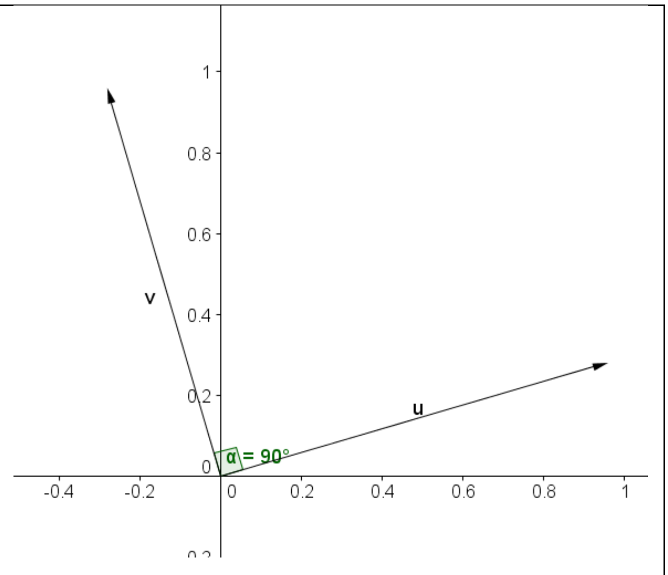
which implies that

$$\boldsymbol{A}^{-1} = \boldsymbol{A}^\top\,, \tag{3.30}$$

i.e., the inverse is obtained by simply transposing the matrix.

Example:

$$A = \begin{bmatrix} 0.96 & -0.28 \\ 0.28 & 0.96 \end{bmatrix}$$



Transformations by orthogonal matrices are special because the length of a vector $\boldsymbol{x}$ is not changed when transforming it using an orthogonal matrix $\boldsymbol{A}$. For the dot product, we obtain

$$\|\boldsymbol{A}\boldsymbol{x}\|^2 = (\boldsymbol{A}\boldsymbol{x})^\top(\boldsymbol{A}\boldsymbol{x}) = \boldsymbol{x}^\top\boldsymbol{A}^\top\boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^\top\boldsymbol{I}\boldsymbol{x} = \boldsymbol{x}^\top\boldsymbol{x} = \|\boldsymbol{x}\|^2\,. \tag{3.31}$$

Moreover, the angle between any two vectors $\boldsymbol{x}, \boldsymbol{y}$, as measured by their inner product, is also unchanged when transforming both of them using an orthogonal matrix $\boldsymbol{A}$. Assuming the dot product as the inner product, the angle of the images $\boldsymbol{A}\boldsymbol{x}$ and $\boldsymbol{A}\boldsymbol{y}$ is given as

$$\cos\omega = \frac{(\boldsymbol{A}\boldsymbol{x})^\top(\boldsymbol{A}\boldsymbol{y})}{\|\boldsymbol{A}\boldsymbol{x}\|\,\|\boldsymbol{A}\boldsymbol{y}\|} = \frac{\boldsymbol{x}^\top\boldsymbol{A}^\top\boldsymbol{A}\boldsymbol{y}}{\sqrt{\boldsymbol{x}^\top\boldsymbol{A}^\top\boldsymbol{A}\boldsymbol{x}\boldsymbol{y}^\top\boldsymbol{A}^\top\boldsymbol{A}\boldsymbol{y}}} = \frac{\boldsymbol{x}^\top\boldsymbol{y}}{\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|}\,, \tag{3.32}$$

1.2.3 Orthonormal Basis

Definition 3.9 (Orthonormal Basis). Consider an n-dimensional vector space V and a basis $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ of V. If

$$\langle \boldsymbol{b}_i, \boldsymbol{b}_j \rangle = 0 \quad \text{for } i \neq j \tag{3.33}$$

$$\langle \boldsymbol{b}_i, \boldsymbol{b}_i \rangle = 1 \tag{3.34}$$

for all $i, j = 1, \ldots, n$ then the basis is called an *orthonormal basis* (*ONB*). If only (3.33) is satisfied, then the basis is called an *orthogonal basis*. Note that (3.34) implies that every basis vector has length/norm 1.

Example 3.8 (Orthonormal Basis)
The canonical/standard basis for a Euclidean vector space $\mathbb{R}^n$ is an orthonormal basis, where the inner product is the dot product of vectors.
In $\mathbb{R}^2$, the vectors

$$\boldsymbol{b}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}, \quad \boldsymbol{b}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix} \tag{3.35}$$

form an orthonormal basis since $\boldsymbol{b}_1^\top \boldsymbol{b}_2 = 0$ and $\|\boldsymbol{b}_1\| = 1 = \|\boldsymbol{b}_2\|$.

1.2.4 Orthogonal Complement

The orthogonal complement of a vector subspace, S , of V, denoted by $S^\perp$, is the set of all vectors that are orthogonal to all elements of S, that is:

$$S^\perp = \{x \in V : \langle x, y \rangle = 0, \text{ for all } y \in S\}.$$

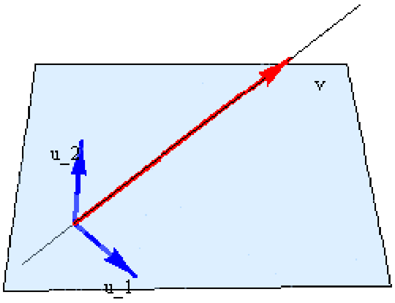
Theorem

Let $U \subseteq V$ a subspace of V such that $S = span[v_1, v_2, \ldots . v_p]$, then

- $v \in S^\perp$ if and only if $< v, v_i >= 0$, forall v_i.
- $S^\perp$ is a subspace
- $\dim(S^\perp) + \dim(S) = n$.
- $v = z_1 + z_2$, with $z_1 \in S$ and $z_2 \in S^\perp$

Consider the plane in $\mathbb{R}^3$ $S = \left\{\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 : -x - 4y + 7z = 0\right\}$, $B = \left\{\begin{bmatrix} -4 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 7 \\ 0 \\ 1 \end{bmatrix}\right\}$ a basis of.

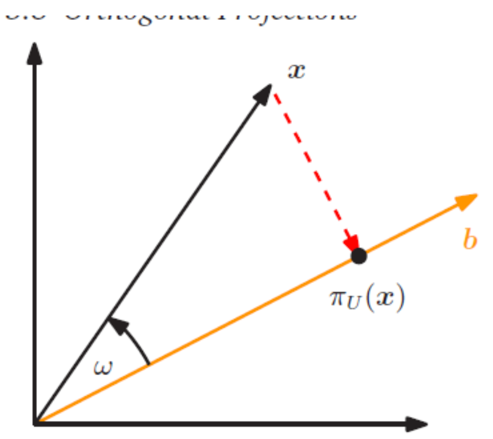
$$v = \begin{bmatrix} 1 \\ 4 \\ -7 \end{bmatrix} \in S^{\perp}$$



1.3 Orthogonal Projections

Definition 3.10 (Projection). Let V be a vector space and $U \subseteq V$ a subspace of V. A linear mapping $\pi : V \to U$ is called a *projection* if $\pi^2 = \pi \circ \pi = \pi$.

1.3.1 Projection onto One-Dimensional Subspaces (Lines)



(a) Projection of $x \in \mathbb{R}^2$ onto a subspace U with basis vector b.

1. Finding the coordinate λ. The orthogonality condition yields

$$\langle \boldsymbol{x} - \pi_U(\boldsymbol{x}), \boldsymbol{b} \rangle = 0 \overset{\pi_U(\boldsymbol{x})=\lambda \boldsymbol{b}}{\iff} \langle \boldsymbol{x} - \lambda \boldsymbol{b}, \boldsymbol{b} \rangle = 0 \,. \tag{3.39}$$

We can now exploit the bilinearity of the inner product and arrive at

$$\langle \boldsymbol{x}, \boldsymbol{b} \rangle - \lambda \langle \boldsymbol{b}, \boldsymbol{b} \rangle = 0 \iff \lambda = \frac{\langle \boldsymbol{x}, \boldsymbol{b} \rangle}{\langle \boldsymbol{b}, \boldsymbol{b} \rangle} = \frac{\langle \boldsymbol{b}, \boldsymbol{x} \rangle}{\|\boldsymbol{b}\|^2} \,. \tag{3.40}$$

In the last step, we exploited the fact that inner products are symmetric. If we choose $\langle \cdot, \cdot \rangle$ to be the dot product, we obtain

$$\lambda = \frac{\boldsymbol{b}^\top \boldsymbol{x}}{\boldsymbol{b}^\top \boldsymbol{b}} = \frac{\boldsymbol{b}^\top \boldsymbol{x}}{\|\boldsymbol{b}\|^2} \,. \tag{3.41}$$

If $\|\boldsymbol{b}\| = 1$, then the coordinate λ of the projection is given by $\boldsymbol{b}^\top \boldsymbol{x}$.

2. Finding the projection point $\pi_U(\boldsymbol{x}) \in U$. Since $\pi_U(\boldsymbol{x}) = \lambda \boldsymbol{b}$, we immediately obtain with (3.40) that

$$\pi_U(\boldsymbol{x}) = \lambda \boldsymbol{b} = \frac{\langle \boldsymbol{x}, \boldsymbol{b} \rangle}{\|\boldsymbol{b}\|^2} \boldsymbol{b} = \frac{\boldsymbol{b}^\top \boldsymbol{x}}{\|\boldsymbol{b}\|^2} \boldsymbol{b} \,, \tag{3.42}$$

where the last equality holds for the dot product only. We can also compute the length of $\pi_U(\boldsymbol{x})$ by means of Definition 3.1 as

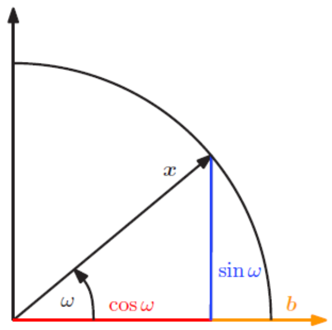
$$\|\pi_U(\boldsymbol{x})\| = \|\lambda \boldsymbol{b}\| = |\lambda| \, \|\boldsymbol{b}\| \,. \tag{3.43}$$

Hence, our projection is of length $|\lambda|$ times the length of $\boldsymbol{b}$. This also adds the intuition that λ is the coordinate of $\pi_U(\boldsymbol{x})$ with respect to the basis vector $\boldsymbol{b}$ that spans our one-dimensional subspace U.

If we use the dot product as an inner product, we get

$$\|\pi_U(\boldsymbol{x})\| \overset{(3.42)}{=} \frac{|\boldsymbol{b}^\top \boldsymbol{x}|}{\|\boldsymbol{b}\|^2} \|\boldsymbol{b}\| \overset{(3.25)}{=} |\cos \omega| \, \|\boldsymbol{x}\| \, \|\boldsymbol{b}\| \frac{\|\boldsymbol{b}\|}{\|\boldsymbol{b}\|^2} = |\cos \omega| \, \|\boldsymbol{x}\| \,. \tag{3.44}$$

83



(b) Projection of a two-dimensional vector $\boldsymbol{x}$ with $\|\boldsymbol{x}\| = 1$ onto a one-dimensional subspace spanned by $\boldsymbol{b}$.

illustration is given in Figure 3.10(b).

3. Finding the projection matrix $\boldsymbol{P}_\pi$. We know that a projection is a linear mapping (see Definition 3.10). Therefore, there exists a projection matrix $\boldsymbol{P}_\pi$, such that $\pi_U(\boldsymbol{x}) = \boldsymbol{P}_\pi \boldsymbol{x}$. With the dot product as inner product and

$$\pi_U(\boldsymbol{x}) = \lambda \boldsymbol{b} = \boldsymbol{b}\lambda = \boldsymbol{b}\frac{\boldsymbol{b}^\top \boldsymbol{x}}{\|\boldsymbol{b}\|^2} = \frac{\boldsymbol{b}\boldsymbol{b}^\top}{\|\boldsymbol{b}\|^2}\boldsymbol{x}\,, \tag{3.45}$$

we immediately see that

$$\boldsymbol{P}_\pi = \frac{\boldsymbol{b}\boldsymbol{b}^\top}{\|\boldsymbol{b}\|^2}\,. \tag{3.46}$$

Note that $\boldsymbol{b}\boldsymbol{b}^\top$ (and, consequently, $\boldsymbol{P}_\pi$) is a symmetric matrix (of rank 1), and $\|\boldsymbol{b}\|^2 = \langle \boldsymbol{b}, \boldsymbol{b}\rangle$ is a scalar.

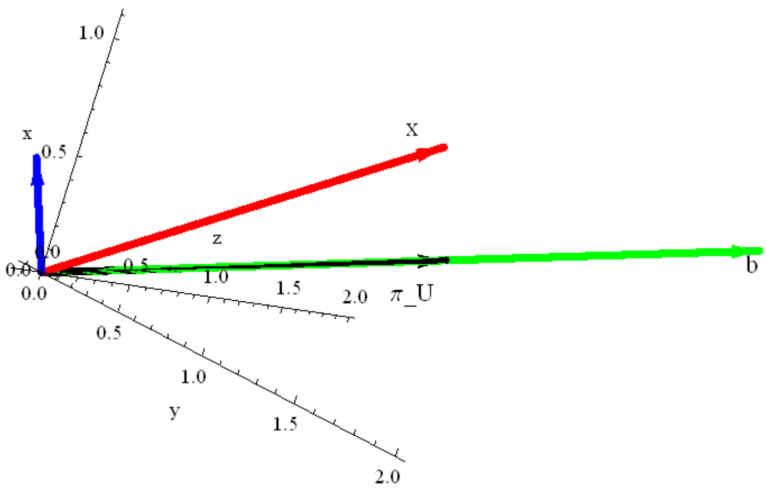
Example 3.10 (Projection onto a Line)
Find the projection matrix $\boldsymbol{P}_\pi$ onto the line through the origin spanned by $\boldsymbol{b} = \begin{bmatrix}1 & 2 & 2\end{bmatrix}^\top$. $\boldsymbol{b}$ is a direction and a basis of the one-dimensional subspace (line through origin).

With (3.46), we obtain

$$\boldsymbol{P}_\pi = \frac{\boldsymbol{b}\boldsymbol{b}^\top}{\boldsymbol{b}^\top\boldsymbol{b}} = \frac{1}{9}\begin{bmatrix}1\\2\\2\end{bmatrix}\begin{bmatrix}1 & 2 & 2\end{bmatrix} = \frac{1}{9}\begin{bmatrix}1 & 2 & 2\\2 & 4 & 4\\2 & 4 & 4\end{bmatrix}. \tag{3.47}$$

Let us now choose a particular $\boldsymbol{x}$ and see whether it lies in the subspace spanned by $\boldsymbol{b}$. For $\boldsymbol{x} = \begin{bmatrix}1 & 1 & 1\end{bmatrix}^\top$, the projection is

$$\pi_U(\boldsymbol{x}) = \boldsymbol{P}_\pi \boldsymbol{x} = \frac{1}{9}\begin{bmatrix}1 & 2 & 2\\2 & 4 & 4\\2 & 4 & 4\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix} = \frac{1}{9}\begin{bmatrix}5\\10\\10\end{bmatrix} \in \operatorname{span}[\begin{bmatrix}1\\2\\2\end{bmatrix}]. \tag{3.48}$$



1.3.2 Projection onto General Subspaces

- $x \in \mathbb{R}^n$, $U \subseteq \mathbb{R}$, $\dim(U) = m \geq 1$
- $(b_1, b_2, \dots, b_m)$ an ordered basis of U

1. Find the coordinates $\lambda_1, \dots, \lambda_m$ of the projection (with respect to the basis of U), such that the linear combination

$$\pi_U(\boldsymbol{x}) = \sum_{i=1}^{m} \lambda_i \boldsymbol{b}_i = \boldsymbol{B\lambda}\,, \tag{3.49}$$

$$\boldsymbol{B} = [\boldsymbol{b}_1, \dots, \boldsymbol{b}_m] \in \mathbb{R}^{n \times m}, \quad \boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_m]^\top \in \mathbb{R}^m\,, \tag{3.50}$$

is closest to $\boldsymbol{x} \in \mathbb{R}^n$. As in the 1D case, "closest" means "minimum distance", which implies that the vector connecting $\pi_U(\boldsymbol{x}) \in U$ and $\boldsymbol{x} \in \mathbb{R}^n$ must be orthogonal to all basis vectors of U. Therefore, we obtain m simultaneous conditions (assuming the dot product as the inner product)

$$\langle \boldsymbol{b}_1, \boldsymbol{x} - \pi_U(\boldsymbol{x}) \rangle = \boldsymbol{b}_1^\top (\boldsymbol{x} - \pi_U(\boldsymbol{x})) = 0 \tag{3.51}$$

$$\vdots$$

$$\langle \boldsymbol{b}_m, \boldsymbol{x} - \pi_U(\boldsymbol{x}) \rangle = \boldsymbol{b}_m^\top (\boldsymbol{x} - \pi_U(\boldsymbol{x})) = 0 \tag{3.52}$$

which, with $\pi_U(\boldsymbol{x}) = \boldsymbol{B\lambda}$, can be written as

$$\boldsymbol{b}_1^\top (\boldsymbol{x} - \boldsymbol{B\lambda}) = 0 \tag{3.53}$$

$$\vdots$$

$$\boldsymbol{b}_m^\top (\boldsymbol{x} - \boldsymbol{B\lambda}) = 0 \tag{3.54}$$

such that we obtain a homogeneous linear equation system

$$\begin{bmatrix} \boldsymbol{b}_1^\top \\ \vdots \\ \boldsymbol{b}_m^\top \end{bmatrix} \begin{bmatrix} \boldsymbol{x} - \boldsymbol{B\lambda} \end{bmatrix} = \boldsymbol{0} \iff \boldsymbol{B}^\top (\boldsymbol{x} - \boldsymbol{B\lambda}) = \boldsymbol{0} \tag{3.55}$$

$$\iff \boldsymbol{B}^\top \boldsymbol{B\lambda} = \boldsymbol{B}^\top \boldsymbol{x}\,. \tag{3.56}$$

The last expression is called *normal equation*. Since $\boldsymbol{b}_1, \dots, \boldsymbol{b}_m$ are a basis of U and, therefore, linearly independent, $\boldsymbol{B}^\top \boldsymbol{B} \in \mathbb{R}^{m \times m}$ is regular and can be inverted. This allows us to solve for the coefficients/ coordinates

$$\boldsymbol{\lambda} = (\boldsymbol{B}^\top \boldsymbol{B})^{-1} \boldsymbol{B}^\top \boldsymbol{x}\,. \tag{3.57}$$

The matrix $(\boldsymbol{B}^\top \boldsymbol{B})^{-1} \boldsymbol{B}^\top$ is also called the *pseudo-inverse* of $\boldsymbol{B}$, which can be computed for non-square matrices $\boldsymbol{B}$. It only requires that $\boldsymbol{B}^\top \boldsymbol{B}$ is positive definite, which is the case if $\boldsymbol{B}$ is full rank. In practical applications (e.g., linear regression), we often add a "jitter term" $\epsilon \boldsymbol{I}$ to

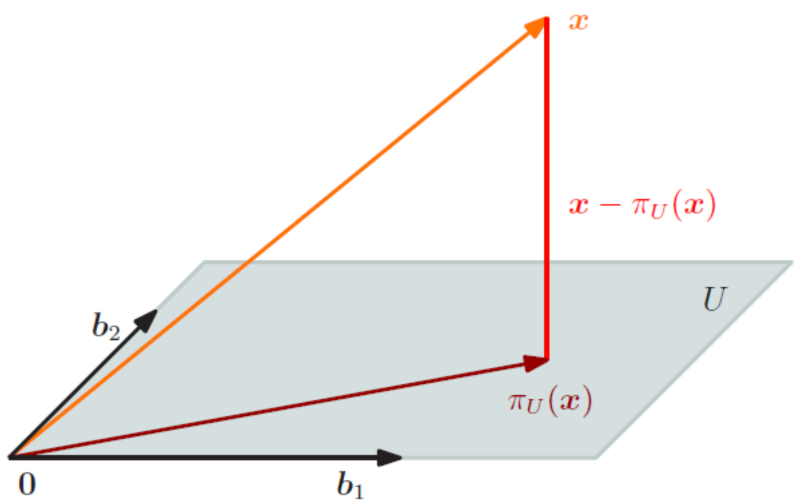
2. Find the projection $\pi_U(\boldsymbol{x}) \in U$. We already established that $\pi_U(\boldsymbol{x}) = \boldsymbol{B\lambda}$. Therefore, with (3.57)

$$\pi_U(\boldsymbol{x}) = \boldsymbol{B}(\boldsymbol{B}^\top\boldsymbol{B})^{-1}\boldsymbol{B}^\top\boldsymbol{x}\,. \tag{3.58}$$

3. Find the projection matrix $\boldsymbol{P}_\pi$. From (3.58), we can immediately see that the projection matrix that solves $\boldsymbol{P}_\pi\boldsymbol{x} = \pi_U(\boldsymbol{x})$ must be

$$\boldsymbol{P}_\pi = \boldsymbol{B}(\boldsymbol{B}^\top\boldsymbol{B})^{-1}\boldsymbol{B}^\top\,. \tag{3.59}$$

Remark. The solution for projecting onto general subspaces includes the 1D case as a special case: If $\dim(U) = 1$, then $\boldsymbol{B}^\top\boldsymbol{B} \in \mathbb{R}$ is a scalar and we can rewrite the projection matrix in (3.59) $\boldsymbol{P}_\pi = \boldsymbol{B}(\boldsymbol{B}^\top\boldsymbol{B})^{-1}\boldsymbol{B}^\top$ as $\boldsymbol{P}_\pi = \frac{\boldsymbol{B}\boldsymbol{B}^\top}{\boldsymbol{B}^\top\boldsymbol{B}}$, which is exactly the projection matrix in (3.46). $\diamondsuit$



Example 3.11 (Projection onto a Two-dimensional Subspace)

For a subspace $U = \text{span}[\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}0\\1\\2\end{bmatrix}] \subseteq \mathbb{R}^3$ and $\boldsymbol{x} = \begin{bmatrix}6\\0\\0\end{bmatrix} \in \mathbb{R}^3$ find the coordinates $\boldsymbol{\lambda}$ of $\boldsymbol{x}$ in terms of the subspace U, the projection point $\pi_U(\boldsymbol{x})$ and the projection matrix $\boldsymbol{P}_\pi$.

First, we see that the generating set of U is a basis (linear independence) and write the basis vectors of U into a matrix $\boldsymbol{B} = \begin{bmatrix}1 & 0\\1 & 1\\1 & 2\end{bmatrix}$.

Second, we compute the matrix $\boldsymbol{B}^\top \boldsymbol{B}$ and the vector $\boldsymbol{B}^\top \boldsymbol{x}$ as

$$\boldsymbol{B}^\top \boldsymbol{B} = \begin{bmatrix}1 & 1 & 1\\0 & 1 & 2\end{bmatrix}\begin{bmatrix}1 & 0\\1 & 1\\1 & 2\end{bmatrix} = \begin{bmatrix}3 & 3\\3 & 5\end{bmatrix}, \quad \boldsymbol{B}^\top \boldsymbol{x} = \begin{bmatrix}1 & 1 & 1\\0 & 1 & 2\end{bmatrix}\begin{bmatrix}6\\0\\0\end{bmatrix} = \begin{bmatrix}6\\0\end{bmatrix}. \tag{3.60}$$

Third, we solve the normal equation $\boldsymbol{B}^\top \boldsymbol{B}\boldsymbol{\lambda} = \boldsymbol{B}^\top \boldsymbol{x}$ to find $\boldsymbol{\lambda}$:

$$\begin{bmatrix}3 & 3\\3 & 5\end{bmatrix}\begin{bmatrix}\lambda_1\\\lambda_2\end{bmatrix} = \begin{bmatrix}6\\0\end{bmatrix} \iff \boldsymbol{\lambda} = \begin{bmatrix}5\\-3\end{bmatrix}. \tag{3.61}$$

Fourth, the projection $\pi_U(\boldsymbol{x})$ of $\boldsymbol{x}$ onto U, i.e., into the column space of $\boldsymbol{B}$, can be directly computed via

$$\pi_U(\boldsymbol{x}) = \boldsymbol{B}\boldsymbol{\lambda} = \begin{bmatrix}5\\2\\-1\end{bmatrix}. \tag{3.62}$$

The corresponding *projection error* is the norm of the difference vector between the original vector and its projection onto U, i.e.,

$$\|\boldsymbol{x} - \pi_U(\boldsymbol{x})\| = \left\|\begin{bmatrix}1 & -2 & 1\end{bmatrix}^\top\right\| = \sqrt{6}. \tag{3.63}$$

Fifth, the projection matrix (for any $\boldsymbol{x} \in \mathbb{R}^3$) is given by

$$\boldsymbol{P}_\pi = \boldsymbol{B}(\boldsymbol{B}^\top \boldsymbol{B})^{-1}\boldsymbol{B}^\top = \frac{1}{6}\begin{bmatrix}5 & 2 & -1\\2 & 2 & 2\\-1 & 2 & 5\end{bmatrix}. \tag{3.64}$$

To verify the results, we can (a) check whether the displacement vector $\pi_U(\boldsymbol{x}) - \boldsymbol{x}$ is orthogonal to all basis vectors of U, and (b) verify that $\boldsymbol{P}_\pi = \boldsymbol{P}_\pi^2$ (see Definition 3.10).

Remark. The projections $\pi_U(\boldsymbol{x})$ are still vectors in $\mathbb{R}^n$ although they lie in an m-dimensional subspace $U \subseteq \mathbb{R}^n$. However, to represent a projected vector we only need the m coordinates $\lambda_1, \dots, \lambda_m$ with respect to the basis vectors $\boldsymbol{b}_1, \dots, \boldsymbol{b}_m$ of U. $\diamondsuit$

Remark. We just looked at projections of vectors $\boldsymbol{x}$ onto a subspace U with basis vectors $\{\boldsymbol{b}_1, \dots, \boldsymbol{b}_k\}$. If this basis is an ONB, i.e., (3.33) and (3.34) are satisfied, the projection equation (3.58) simplifies greatly to

$$\pi_U(\boldsymbol{x}) = \boldsymbol{B}\boldsymbol{B}^\top \boldsymbol{x} \tag{3.65}$$

since $\boldsymbol{B}^\top \boldsymbol{B} = \boldsymbol{I}$ with coordinates

$$\boldsymbol{\lambda} = \boldsymbol{B}^\top \boldsymbol{x} \,. \tag{3.66}$$

This means that we no longer have to compute the inverse from (3.58), which saves computation time. $\diamondsuit$

1.3.3 Gram-Schmidt Orthogonalization

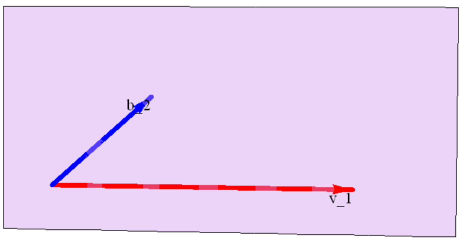
$(b_1, b_2, \dots b_n)$ a basis of U

- 1^{st} *vector*

$$v_1 = \frac{b_1}{||b_2||}$$

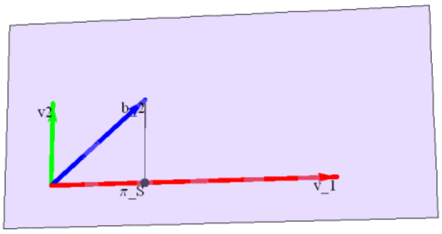
- 2^{nd} *vector*

  $S = span[v_1]$ and $B = [v_1]$



$$v'_2 = b_2 - \pi_S(b_2) = b_2 - BB^\top b_2$$

$$v_2 = \frac{b'_2}{||b'_2||}$$

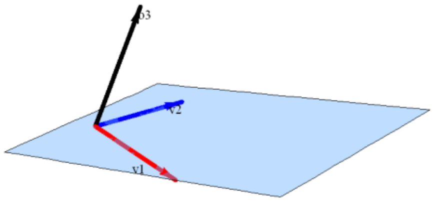


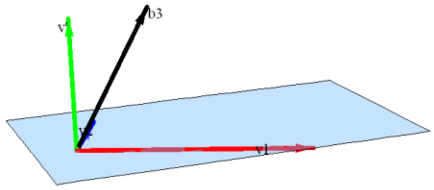
- $3^{th}\ vector$

  $S = span[v_1, v2]$ and $B = [v_1\ v_2]$

$v'_3 = b_3 - \pi_S(b_3) = b_3 - BB^\top b_3$

$v_3 = \frac{v'_3}{||v'_3||}$





- $p^{th}\ \ vector$

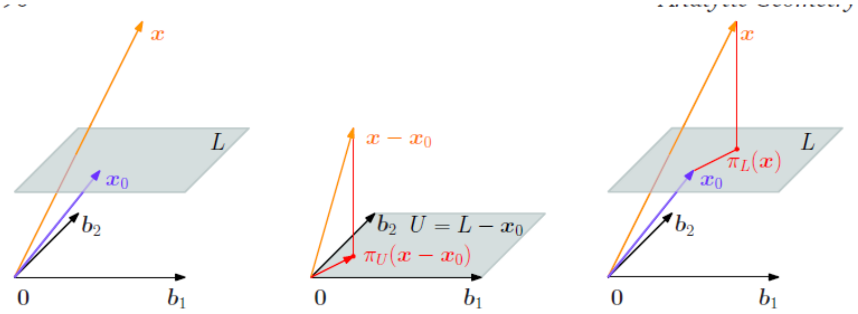
  $S = span[v_1, v2, \ldots v_{p-1}]$ and $B = [v_1\ v_2\ \ldots.\ v_{p-1}]$

$v'_p = b_p - \pi_S(b_p) = b_p - BB^\top b_p$

$v_p = \frac{v'_p}{||v'_p||}$

1.3.4 Projection onto Affine Subspaces

- $L = x_o + U$ affine subspace.
- (b_1, b_2) a basis of U.



(a) Setting.

(b) Reduce problem to projection π_U onto vector subspace.

(c) Add support point back in to get affine projection π_L.

To determine the orthogonal projection, π_L of x onto L, we transform the problem into a problem that we know how to solve: the projection onto a vector subspace.

$L - x_o = U$, then we use the orthogonal projection $\pi(x - x_0)$

$$\pi_L(\boldsymbol{x}) = \boldsymbol{x}_0 + \pi_U(\boldsymbol{x} - \boldsymbol{x}_0)\,,$$

$$\begin{aligned} d(\boldsymbol{x}, L) &= \|\boldsymbol{x} - \pi_L(\boldsymbol{x})\| = \|\boldsymbol{x} - (\boldsymbol{x}_0 + \pi_U(\boldsymbol{x} - \boldsymbol{x}_0))\| \\ &= d(\boldsymbol{x} - \boldsymbol{x}_0, \pi_U(\boldsymbol{x} - \boldsymbol{x}_0))\,. \end{aligned}$$