A norm for matrices

Once we can measure the size of a vector $\mathbf{v}$ by a norm $\|\mathbf{v}\|$, we may think about measuring the "size" of a linear map. Consider $A \in \mathbb{R}^{m\times n}$, and $\mathbf{w} = A\mathbf{v}$. Then the following quotient

$$\frac{\|\mathbf{w}\|_{\mathbb{R}^m}}{\|\mathbf{v}\|_{\mathbb{R}^n}} = \frac{\|A\mathbf{v}\|_{\mathbb{R}^m}}{\|\mathbf{v}\|_{\mathbb{R}^n}}$$

tells us, how much longer (or shorter) $\mathbf{w} = A\mathbf{v}$ is, compared to $\mathbf{v}$. A should be "large", if it produces long vectors from short ones, and "small", if it produces short vectors from long ones. Thus, we may define

$$\|A\| := \max_{\mathbf{v}\neq 0} \frac{\|A\mathbf{v}\|_{\mathbb{R}^m}}{\|\mathbf{v}\|_{\mathbb{R}^n}},$$

so that we have:

$$\|\mathbf{w}\|_{\mathbb{R}^m} = \|A\mathbf{v}\|_{\mathbb{R}^m} \le \|A\|\|\mathbf{v}\|_{\mathbb{R}^n}.$$

It is not easy to compute this norm. We will consider a possibility later.

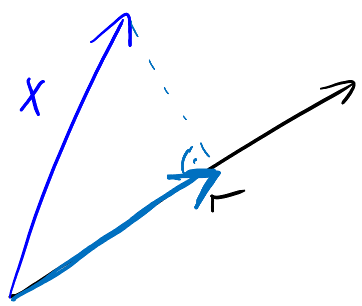
VL12 ↓

5.2 Orthogonal projections

In this section $\langle\cdot,\cdot\rangle$ denotes an arbitrary inner product in $\mathbb{R}^n$.

5.2.1 Orthogonal projection onto a line

Imagine you ride a rowboat on a river. You want to go in a direction $\mathbf{r} \neq \mathbf{o}$. However water flows in direction $\mathbf{x}$, which is not parallel to $\mathbf{r}$.



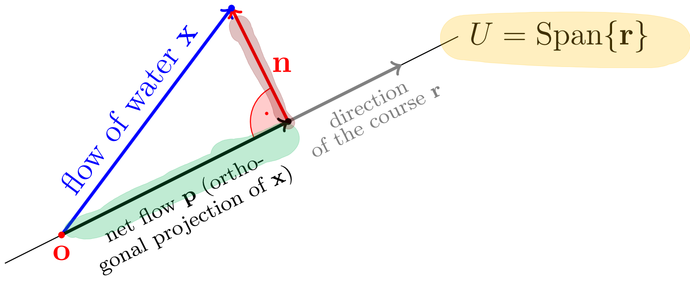
The steersman asks:

What is the component of $\mathbf{x}$ *with respect to the wanted direction* $\mathbf{r}$*?*

In mathematical language: Write the vector $\mathbf{x}$ in a linear combination

$$\mathbf{x} = \mathbf{p} + \mathbf{n}$$

consisting of two orthogonal vectors: $\mathbf{p}$ is parallel to the wanted direction $\mathbf{r}$ and $\mathbf{n}$ is orthogonal to this.



Definition 5.6. Orthogonal projection onto a line

Let $\langle\cdot,\cdot\rangle$ *be an inner product in* $\mathbb{R}^n$ *and* $U := \mathrm{Span}(\mathbf{r})$ *for* $\mathbf{r} \neq \mathbf{o}$*. For a decomposition* $\mathbf{x} = \mathbf{p} + \mathbf{n}$ *for a vector* $\mathbf{x} \in \mathbb{R}^n$ *into two orthogonal vectors* $\mathbf{p} \in U$ *and* $\mathbf{n} \perp \mathbf{r}$*, we call* $\mathbf{p}$ *the orthogonal projection of* $\mathbf{x}$ *onto* U*, and* $\mathbf{n}$ *is called the normal component of* $\mathbf{x}$ *with respect to* U*.*

Show: There is only one such decomposition.

$x = p + n \quad , \quad x = p' + n' \quad$ with $\quad p, p' \in U \quad , \quad n, n' \perp r \quad (n, n' \in U^\perp)$.

$$\Rightarrow \underset{\in U}{p - p'} = \underset{\substack{(\in U^\perp) \\ n'-n \perp r}}{n' - n} \qquad \Rightarrow \begin{cases} \langle p-p', p-p'\rangle = \langle p-p', n'-n\rangle = 0 \\ \underbrace{\langle n'-n, p-p'\rangle}_{=0} = \langle n'-n, n'-n\rangle \end{cases}$$

$$\overset{(S1)}{\Rightarrow} p - p' = 0 \quad , \quad n' - n = 0 \quad \Rightarrow \quad p = p' \, , \, n' = n .$$

Calculation of p and n: Because of $\mathbf{p} \in U = \mathrm{Span}(\mathbf{r})$, we have $\mathbf{p} = \lambda\mathbf{r}$ for a $\lambda \in \mathbb{R}$, which we simply have to find. Since $\mathbf{x} = \mathbf{p} + \mathbf{n} = \lambda\mathbf{r} + \mathbf{n}$ and $\mathbf{n} \perp \mathbf{r}$, we get:

$$\langle \mathbf{x}, \mathbf{r}\rangle = \langle p+n, r\rangle = \langle \lambda r + n, r\rangle \overset{(S2)}{=} \langle \lambda r, r\rangle + \underbrace{\langle n, r\rangle}_{=0} \overset{(S3)}{=} \lambda \langle r, r\rangle$$

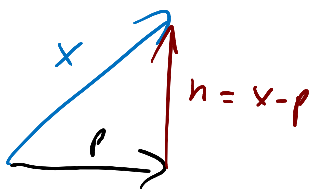
$$\Rightarrow \lambda = \frac{\langle x, r\rangle}{\langle r, r\rangle}$$

The case $\mathbf{r} = \mathbf{o}$ (i.e. $\mathbf{p} = \mathbf{o}$ and $\mathbf{n} = \mathbf{x}$) is omitted here. In summary, we get:

Proposition 5.7. Orth. projection & and normal component w.r.t a line

Let $\mathbf{x}, \mathbf{r} \in \mathbb{R}^n$ *with* $\mathbf{r} \neq \mathbf{o}$*. For the orthogonal projection* $\mathbf{p}$ *of* $\mathbf{x}$ *onto* $U = \mathrm{Span}(\mathbf{r})$ *and the associated normal component* $\mathbf{n}$ *of* $\mathbf{x}$ *w.r.t.* U*, one finds:*

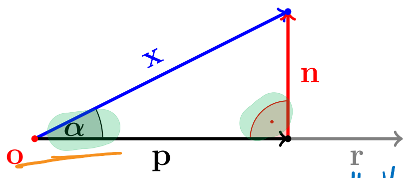
$$\mathbf{p} = \frac{\langle \mathbf{x}, \mathbf{r}\rangle}{\langle \mathbf{r}, \mathbf{r}\rangle}\mathbf{r} \qquad \textit{and} \qquad \mathbf{n} = \mathbf{x} - \mathbf{p} = \mathbf{x} - \frac{\langle \mathbf{x}, \mathbf{r}\rangle}{\langle \mathbf{r}, \mathbf{r}\rangle}\mathbf{r}.$$



Rule of thumb: $\|\cdot\|$ gives length and $\langle\cdot,\cdot\rangle$ gives an angle

*Geometrically $\|\mathbf{x}\|$ is seen as a **length** of the vector $\mathbf{x}$. The inner product $\langle \mathbf{x}, \mathbf{y}\rangle$ gives back the **angle** between $\mathbf{x}$ and $\mathbf{y}$.*

To define a meaningful angle between vectors, we again look at the triangle, given by the vectors $\mathbf{x}$, $\mathbf{p}$ and $\mathbf{n}$. It is right-angled since $\mathbf{p} \perp \mathbf{n}$ is our definition of 90 degree. The angle between $\mathbf{x}$ and $\mathbf{r}$ is called α in the picture.



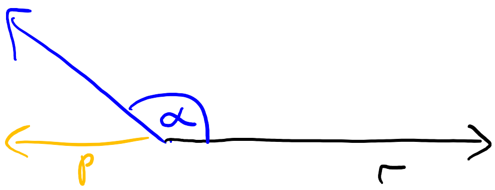
$\cos(\alpha) = \frac{\|p\|}{\|x\|}$

If α is an acute angle, i.e. $\alpha \in [0, \pi/2]$, then $\lambda \geq 0$ and:

$$\|x\| \cos(\alpha) = \|\mathbf{p}\| = \|\lambda r\| = \underbrace{|\lambda|}_{\geq 0} \cdot \|r\| = \frac{\langle x, r\rangle}{\underbrace{\langle r, r\rangle}_{= \|r\|^2}} \cdot \|r\| = \langle x, r\rangle \cdot \frac{1}{\|r\|}$$

We reformulate this:

$$\langle \mathbf{x}, \mathbf{r}\rangle = \|\mathbf{x}\|\|\mathbf{r}\| \cos(\alpha).$$



If α is not acute, we can do an analogue calculation. In summary, we can give the following definition for an angle:

Definition 5.8. Angle between two vectors in $\mathbb{R}^n$

For two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n \setminus \{\mathbf{o}\}$ we write $\operatorname{angle}(\mathbf{x}, \mathbf{y})$ for the angle $\alpha \in [0, \pi]$ between $\mathbf{x}$ and $\mathbf{y}$, which is defined by

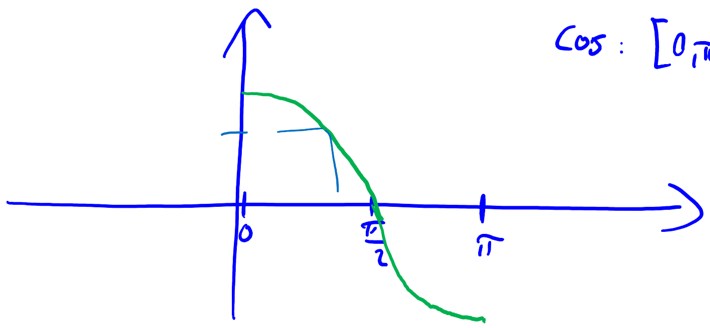
$$\cos(\alpha) = \frac{\langle \mathbf{x}, \mathbf{y}\rangle}{\|\mathbf{x}\|\|\mathbf{y}\|}. \tag{5.3}$$

$\rightsquigarrow \operatorname{angle}(x,y) := \arccos(\text{〃})$

Using Proposition 5.5 (Cauchy-Schwarz-inequality), we conclude that the angle is well-defined:

$$\frac{|\langle \mathbf{x}, \mathbf{y}\rangle|}{\|\mathbf{x}\|\|\mathbf{y}\|} \leq 1 \quad \text{and hence} \quad -1 \leq \frac{\langle \mathbf{x}, \mathbf{y}\rangle}{\|\mathbf{x}\|\|\mathbf{y}\|} \leq 1 .$$

This means the right-hand side of (5.3) is indeed in the range of the cos function. Restricted to $\alpha \in [0, \pi]$, we know that $\cos(\alpha)$ is bijective, and hence, α is well-defined by equation (5.3).



$\cos : [0, \pi] \to [-1, 1]$ bijective

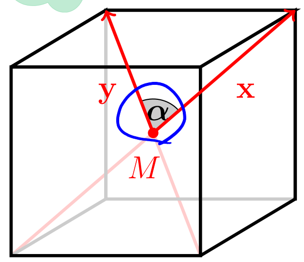
Example 5.9. Consider the cube C in $\mathbb{R}^3$ with center M in the origin and the corners $(\pm1, \pm1, \pm1)^T$, where all the combinations with $\pm$-signs occur.

$\begin{pmatrix}-1\\1\\1\end{pmatrix}$ $\begin{pmatrix}1\\1\\1\end{pmatrix}$

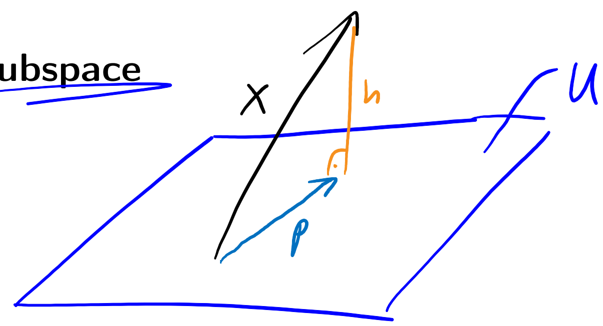
All diagonals of C go trough M and intersect with an angle α, which is calculated with the vectors $\mathbf{x} = (1, 1, 1)^T$ and $\mathbf{y} = (-1, 1, 1)^T$:

$$\cos(\alpha) = \frac{\langle \mathbf{x}, \mathbf{y}\rangle_{euklid}}{\|\mathbf{x}\|\|\mathbf{y}\|} = \frac{-1+1+1}{\sqrt{1+1+1}\sqrt{1+1+1}} = \frac{1}{3},$$

which implies $\alpha = \arccos(\frac{1}{3}) \approx 70.53°$.



5.2.2 Orthogonal projection onto a subspace



$\langle \cdot, \cdot \rangle$ inner product in $\mathbb{R}^n$

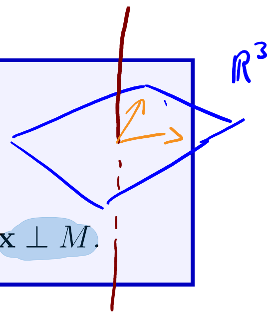
In order to do this, we recall the concept of orthogonal complements:

Definition 5.10. Orthogonal complement $M^\perp$

Let $M \subset \mathbb{R}^n$ be nonempty. Then we call

$$M^\perp := \{\mathbf{x} \in \mathbb{R}^n : \langle \mathbf{x}, \mathbf{m}\rangle = 0 \text{ for all } \mathbf{m} \in M\}$$

the orthogonal complement for M. Instead of $\mathbf{x} \in M^\perp$, we often write $\mathbf{x} \perp M$.

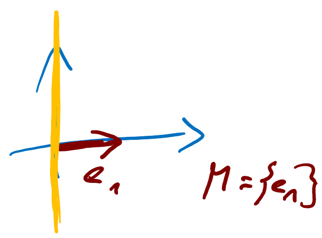


Example 5.11. Consider $\langle \cdot, \cdot\rangle_{euklid}$ the standard inner product in $\mathbb{R}^n$.

(a) For $M = \{\mathbf{o}\}$ in $\mathbb{R}^n$, we have $M^\perp = \mathbb{R}^n$.

$$M^\perp = \{x \in \mathbb{R}^n : \langle x, o\rangle = 0\} = \mathbb{R}^n$$

(b) For $M = \{\mathbf{e}_1\}$ in $\mathbb{R}^2$, we have $M^\perp = \mathrm{Span}(\mathbf{e}_2) \subset \mathbb{R}^2$.

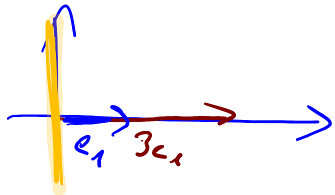


$$M^\perp = \left\{\begin{pmatrix}x_1\\x_2\end{pmatrix} \in \mathbb{R}^2 : \underbrace{\left\langle \begin{pmatrix}x_1\\x_2\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\right\rangle}_{= x_1} = 0\right\} =$$

$$= \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : \quad x_2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \; x_2 \in \mathbb{R} \right\}$$

$$= \left\{ x_2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} : \quad x_2 \in \mathbb{R} \right\} = \underline{\mathrm{Span}(e_2)}$$

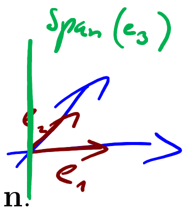
(c) For $M = \{3\mathbf{e}_1\}$ in $\mathbb{R}^2$, we have $M^\perp = \mathrm{Span}(\mathbf{e}_2) \subset \mathbb{R}^2$.



(d) For $M = \{\mathbf{e}_1, 3\mathbf{e}_1\}$ in $\mathbb{R}^2$, we have $M^\perp = \mathrm{Span}(\mathbf{e}_2) \subset \mathbb{R}^2$.

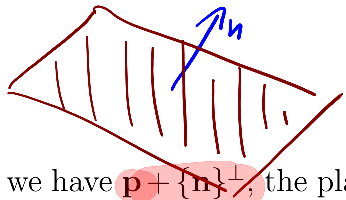
(e) For $M = \mathrm{Span}(\mathbf{e}_1)$ in $\mathbb{R}^2$, we have $M^\perp = \mathrm{Span}(\mathbf{e}_2) \subset \mathbb{R}^2$.

(f) For $M = \{\mathbf{e}_1, \mathbf{e}_2\}$ or $M = \mathrm{Span}(\mathbf{e}_1, \mathbf{e}_2)$ in $\mathbb{R}^3$, we have $M^\perp = \mathrm{Span}(\mathbf{e}_3) \subset \mathbb{R}^3$.

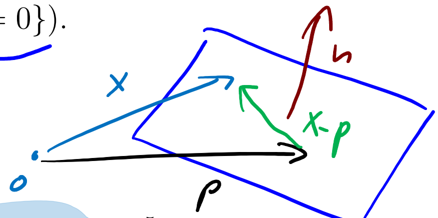


$$M^\perp = \{ \cdots \}$$

(g) For $\mathbf{n} \in \mathbb{R}^3 \setminus \{\mathbf{o}\}$, we have $\{\mathbf{n}\}^\perp$ the plane $\mathbb{R}^3$ through 0 with normal vector $\mathbf{n}$.



(h) For $\mathbf{n} \in \mathbb{R}^3 \setminus \{\mathbf{o}\}$ and $\mathbf{p} \in \mathbb{R}^3$, we have $\mathbf{p} + \{\mathbf{n}\}^\perp$, the plane $\mathbb{R}^3$ through $\mathbf{p}$ with normal vector $\mathbf{n}$ (this is an affine space $\{\mathbf{x} \in \mathbb{R}^3 : \langle \mathbf{x} - \mathbf{p}, \mathbf{n} \rangle = 0\}$).



(i) For $M = \mathrm{Span}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_5)$ in $\mathbb{R}^5$, we have $M^\perp = \mathrm{Span}(\mathbf{e}_3, \mathbf{e}_4) \subset \mathbb{R}^5$.

$$M^\perp = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} \in \mathbb{R}^5 : \left\langle \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ 0 \\ 0 \\ y_5 \end{pmatrix} \right\rangle = 0 \quad \text{for all } y_1, y_2, y_5 \in \mathbb{R} \right\}$$

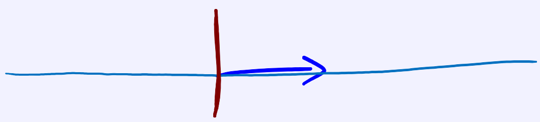
$$= \mathrm{Span}(e_3, e_4)$$

Proposition 5.12.

For all nonempty sets $M \subset \mathbb{R}^n$ we have:

(a) $M^\perp = (\mathrm{Span}(M))^\perp$,

(b) $M^\perp$ is a linear subspace of $\mathbb{R}^n$.



Proof: (a) $M^\perp = (\mathrm{Span}(M))^\perp$. $\quad (\subset)$: $x \in M^\perp \Rightarrow \forall m \in M: \langle x, m \rangle = 0$

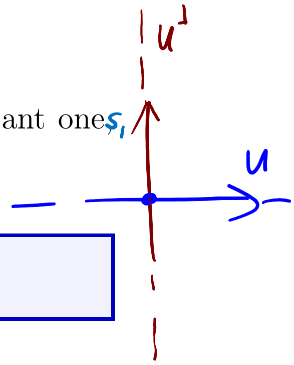
$$\Rightarrow \forall \lambda_j \in \mathbb{R} \;\; \forall m_j \in M: \quad 0 = \sum_{j=1}^{k} \lambda_j \underbrace{\langle x, m_j \rangle}_{=0} \overset{(S2)(S3)}{=} \left\langle x, \sum_{j=1}^{k} \lambda_j m_j \right\rangle$$

$$\Rightarrow \forall u \in \mathrm{Span}(M): \langle x, u \rangle = 0 \Rightarrow x \in (\mathrm{Span}(M))^\perp$$

$(\supset)$ $x \in (\mathrm{Span}(M))^\perp \Rightarrow \forall u \in \mathrm{Span}(M): \langle x, u \rangle = 0 \Rightarrow \forall m \in M: \langle x, m \rangle = 0$

$$\Rightarrow x \in M^\perp$$

(6) Exercise (cf. Homework 4.3) □



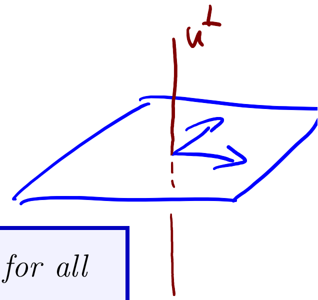
We state one important property of the orthogonal complement. Other important ones, you find at the end of this section.

Proposition 5.13. Properties of $U^\perp$

For a linear subspace $U \subset \mathbb{R}^n$, we have $U \cap U^\perp = \{\mathbf{o}\}$.

Proof: $x \in U \cap U^\perp$. Then $\langle x, u \rangle = 0$ for all $u \in U$, in particular for $u = x$.

$$\Rightarrow \langle x, x \rangle = 0 \underset{(S1)}{\Longrightarrow} x = 0. \quad \square$$



Proposition 5.14. Orthogonal to a basis

Let U be a linear subspace of $\mathbb{R}^n$ and $\mathcal{B} = (\mathbf{u}_1, \ldots, \mathbf{u}_k)$ a basis of U. Then for all $\mathbf{x} \in \mathbb{R}^n$ we have:

$$\mathbf{x} \perp U \quad \Longleftrightarrow \quad \mathbf{x} \perp \mathcal{B}.$$

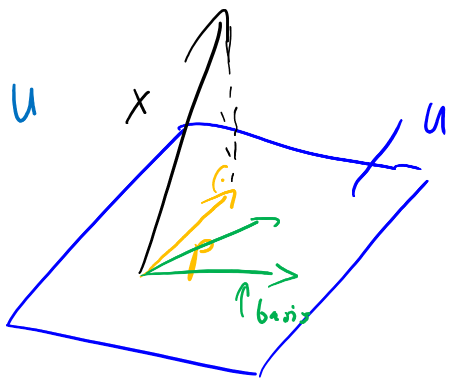
In other words: $\mathbf{x}$ is orthogonal to all vectors in U if and only if it is orthogonal to the basis vectors of U.

Proof: ($\Rightarrow$) ✓ ($\Leftarrow$) $x \perp \mathcal{B} \Rightarrow \langle x, u_j \rangle = 0$ for all j

$$\Rightarrow \forall \lambda_j \in \mathbb{R}: \ 0 = \sum_{j=1}^{k} \lambda_j \underbrace{\langle x, u_j \rangle}_{=0} = \left\langle x, \sum_{j=1}^{k} \lambda_j u_j \right\rangle$$

$$\underset{\mathcal{B}\text{ basis}}{\Longrightarrow} \forall u \in U: \langle x, u \rangle = 0 \Rightarrow x \perp U. \quad \square$$

$\rightsquigarrow$ orth. Projection onto subspace U

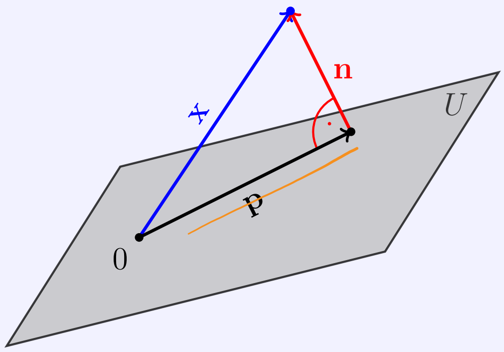


Definition 5.15. Orthogonal projection onto a subspace U

Let U be a linear subspace of $\mathbb{R}^n$ and also let $\mathbf{x} \in \mathbb{R}^n$. Again, we search for a decomposition:

$$\mathbf{p} \in U \quad \text{and} \quad \mathbf{n} \perp U \quad \text{with} \quad \mathbf{x} = \mathbf{p} + \mathbf{n}.$$

In other words, we write $\mathbf{x}$ as a sum of two vectors, where one lies in U and the other one is orthogonal to U.



The (uniquely determined) vector $\mathbf{p}$ is called the orthogonal projection of $\mathbf{x}$ onto U, and $\mathbf{n}$ is called the normal component of $\mathbf{x}$ w.r.t. U.

For the orthogonal projection $\mathbf{p}$ of $\mathbf{x}$ onto U, we often simply write $\mathbf{x}_{|U}$. In this notation, the decomposition $\mathbf{x} = \mathbf{p} + \mathbf{n}$ gets:

$$\mathbf{x} = \mathbf{x}_{|U} + \mathbf{x}_{|U^\perp}, \qquad \text{i.e.} \qquad \mathbf{p} = \mathbf{x}_{|U} \quad \text{and} \quad \mathbf{n} = \mathbf{x}_{|U^\perp}$$

Calculation of the orthogonal projection $\mathbf{x}_{|U}$: Choose a basis of U : $\mathcal{B} = (u_1, \ldots, u_k)$

Then: $p = \alpha_1 u_1 + \cdots + \alpha_k u_k$ with some $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ (Find α_i!)

For $i = 1, \ldots, k$, we have for $x = p + n = \alpha_1 u_1 + \cdots + \alpha_k u_k + n$

(S2),(S3)

$$\langle \mathbf{x}, \mathbf{u}_i \rangle = \langle \alpha_1 \mathbf{u}_1 + \ldots + \alpha_k \mathbf{u}_k + \mathbf{n}, \mathbf{u}_i \rangle = \alpha_1 \langle \mathbf{u}_1, \mathbf{u}_i \rangle + \ldots + \alpha_k \langle \mathbf{u}_k, \mathbf{u}_i \rangle + \underbrace{\langle \mathbf{n}, \mathbf{u}_i \rangle}_{0}. \tag{5.4}$$

For i ⇝ an equation

Now we have k equations and k unknowns $\alpha_1, \ldots, \alpha_k$:

Proposition 5.16. Calculating the projection $\mathbf{x}_{|U}$

Let $\mathbf{x} \in \mathbb{R}^n$ and U be a linear subspace of $\mathbb{R}^n$ where $\mathcal{B} = (\mathbf{u}_1, \ldots, \mathbf{u}_k)$ is a basis of U. Then we get the orthogonal projection

$$\mathbf{x}_{|U} = \alpha_1 \mathbf{u}_1 + \ldots + \alpha_k \mathbf{u}_k,$$

where $\alpha_1, \ldots, \alpha_k$ are given by the (unique) solution of the LES:

$$\begin{pmatrix} \langle \mathbf{u}_1, \mathbf{u}_1 \rangle & \langle \mathbf{u}_2, \mathbf{u}_1 \rangle & \ldots & \langle \mathbf{u}_k, \mathbf{u}_1 \rangle \\ \langle \mathbf{u}_1, \mathbf{u}_2 \rangle & \langle \mathbf{u}_2, \mathbf{u}_2 \rangle & \ldots & \langle \mathbf{u}_k, \mathbf{u}_2 \rangle \\ \vdots & \vdots & & \vdots \\ \langle \mathbf{u}_1, \mathbf{u}_k \rangle & \langle \mathbf{u}_2, \mathbf{u}_k \rangle & \ldots & \langle \mathbf{u}_k, \mathbf{u}_k \rangle \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{pmatrix} = \begin{pmatrix} \langle \mathbf{x}, \mathbf{u}_1 \rangle \\ \langle \mathbf{x}, \mathbf{u}_2 \rangle \\ \vdots \\ \langle \mathbf{x}, \mathbf{u}_k \rangle \end{pmatrix}. \tag{5.5}$$

The $(k \times k)$ matrix on the left-hand side is called the Gramian matrix $G(\mathcal{B})$. The normal component $\mathbf{n} = \mathbf{x}_{|U^\perp}$ is then given by $\mathbf{n} = \mathbf{x} - \mathbf{x}_{|U}$.

Proof: Only to show $G(\mathcal{B})$ invertible $\Leftrightarrow \operatorname{Ker}(G(\mathcal{B})) = \{0\}$. Choose $\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} \in \operatorname{Ker}(G(\mathcal{B}))$.

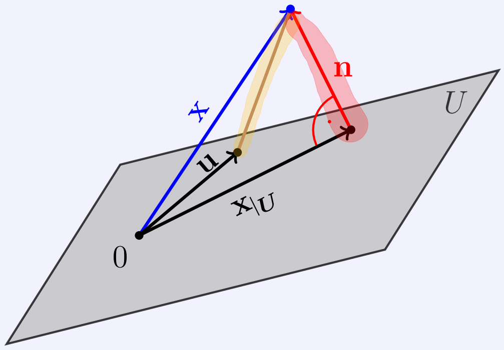
Then: (for all m) $0 = \sum_{i=1}^k \beta_i \langle u_i, u_m \rangle = \langle \underbrace{\overbrace{\sum_{i=1}^k \beta_i u_i}^{=:v}}_{\in U}, \underbrace{u_m}_{\in U} \rangle \Rightarrow v \in U$ and $v \in U^\perp \Rightarrow v = 0$ □

Proposition 5.17. Approximation formula

Let $\mathbf{x} \in \mathbb{R}^n$ and U be a linear subspace of $\mathbb{R}^n$. The orthogonal projection $\mathbf{x}_{|U}$ minimises the distance between $\mathbf{x}$ and the subspace U:

$$\| \underbrace{\mathbf{x} - \mathbf{x}_{|U}}_{\mathbf{n}} \| = \min_{\mathbf{u} \in U} \|\mathbf{x} - \mathbf{u}\| =: \mathrm{dist}(\mathbf{x}, U)$$

In other words: No other vector of U is as closed to $\mathbf{x}$ as $\mathbf{x}_{|U}$.



Proof. For all $\mathbf{u} \in U$, we get

$$\|\mathbf{x} - \mathbf{u}\|^2 = \| (\underbrace{\mathbf{x} - \mathbf{x}_{|U}}_{\mathbf{n}}) + (\underbrace{\mathbf{x}_{|U} - \mathbf{u}}_{=:\mathbf{v}}) \|^2 = \langle \mathbf{n} + \mathbf{v}, \mathbf{n} + \mathbf{v} \rangle = \underbrace{\langle \mathbf{n}, \mathbf{n} \rangle}_{\|\mathbf{n}\|^2} + 2 \underbrace{\langle \mathbf{n}, \mathbf{v} \rangle}_{0} + \underbrace{\langle \mathbf{v}, \mathbf{v} \rangle}_{\geq 0} \geq \|\mathbf{n}\|^2,$$

and, hence, $\|\mathbf{x} - \mathbf{u}\| \geq \|\mathbf{n}\| = \|\mathbf{x} - \mathbf{x}_{|U}\|$. Equality holds if and only if $\mathbf{v} = \mathbf{o}$, i.e. $\mathbf{u} = \mathbf{x}_{|U}$. □

Proposition 5.18.

For all nonempty sets $M \subset \mathbb{R}^n$ we have:

(a) $\mathbb{R}^n = \mathrm{Span}(M) + M^\perp$ *and* $\mathrm{Span}(M) \cap M^\perp = \{\mathbf{o}\}$,

(b) $(M^\perp)^\perp = \mathrm{Span}(M)$.

Proof: (a) Using Prop. 5.16: For $x \in \mathbb{R}^n$: $x = p + n$ with $p \in \mathrm{Span}(M)$, $n \perp \mathrm{Span}(M)$

$\Rightarrow n \in (\mathrm{Span}(M))^\perp = M^\perp$ (Proposition 5.12)

Prop. 5.13 $\Rightarrow \mathrm{Span}(M) \cap M^\perp = \{o\}$

(b) Exercise!

Corollary 5.19. Properties of $U^\perp$

For a linear subspace $U \subset \mathbb{R}^n$, we have:

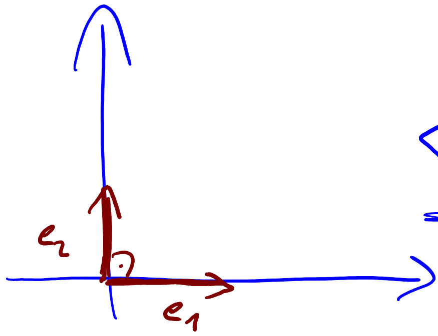
(a) $\mathbb{R}^n = U + U^\perp$ *and* $U \cap U^\perp = \{\mathbf{o}\}$. *Usually, one writes in this case:*

$$\mathbb{R}^n = U \oplus U^\perp.$$

(b) $\dim(U^\perp) = \dim(\mathbb{R}^n) - \dim(U)$.

(c) $(U^\perp)^\perp = U$.

Proof: Exercise!



$$\langle e_i, e_j \rangle_{euklid} = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

5.3 Orthonormal systems and bases

For some applications it it very useful to have a set of vectors $\{\mathbf{u}_1 \ldots \mathbf{u}_k\} \subset \mathbb{R}^n$ which are mutually orthogonal:

$$i \neq j \quad \Rightarrow \quad \mathbf{u}_i \perp \mathbf{u}_j \quad \Leftrightarrow \quad \langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0$$

and have unit norm:

$$\|\mathbf{u}_i\| = \sqrt{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} = 1.$$

Using the Kronecker symbol:

$$\delta_{ij} = \begin{cases} 1 & : \quad i = j \\ 0 & : \quad i \neq j \end{cases}$$

we may write this in short:

$$\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \delta_{ij}.$$

Definition 5.20. OS, ONS, OB, ONB

Let U be a linear subspace of $\mathbb{R}^n$. A family $\mathcal{F} = (\mathbf{u}_1, \ldots, \mathbf{u}_k)$ consisting of vectors from U is called:

- *Orthogonal system (OS) if the vectors in $\mathcal{F}$ are mutually orthogonal: $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0$ for all $i, j \in \{1, \ldots, k\}$ with $i \neq j$;*
- *Orthonormal system (ONS) if $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \delta_{ij}$ for all $i, j \in \{1, \ldots, k\}$;*
- *Orthogonal basis (OB) if it is an OS and a basis of U;*
- *Orthonormal basis (ONB) if it is an ONS and a basis of U.*

If $\mathcal{F}$ is an ONB, then the Gram matrix $G(\mathcal{F})$ is the identity matrix and projections are very easily calculable.

Example 5.21. Let $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_{eukl}$ the standard inner product.

(a) The canonical unit vectors

$$\mathbf{e}_1 = (1, 0, \ldots, 0)^T, \quad \mathbf{e}_2 = (0, 1, 0, \ldots, 0)^T, \quad \ldots, \quad \mathbf{e}_n = (0, \ldots, 0, 1)^T$$