Lecture 5.1: Gram-Schmidt algorithm

Optimization and Computational Linear Algebra for Data Science

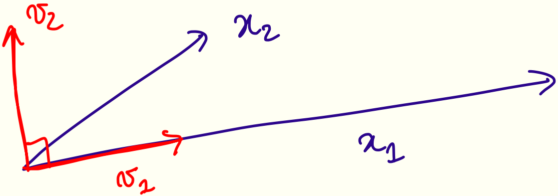
Léo Miolane

Purpose of the algorithm

The Gram-Schmidt process takes as

- **Input:** a *linearly independent* family $(x_1, \ldots, x_k)$ of $\mathbb{R}^n$.
- **Output:** an *orthonormal basis* $(v_1, \ldots v_k)$ of $\mathrm{Span}(x_1, \ldots, x_k)$.

$\|v_1\| = \|v_2\| = 1$



Consequence

Every subspace of $\mathbb{R}^n$ admits an orthonormal basis.

Gram-Schmidt algorithm

The Gram-Schmidt process constructs $v_1, v_2, \ldots, v_k$ in this order, such that for all $i \in \{1, \ldots, k\}$:

$$\mathcal{H}_i : \begin{cases} (v_1, \ldots, v_i) \text{ is an orthonormal family} \\ \mathrm{Span}(v_1, \ldots, v_i) = \mathrm{Span}(x_1, \ldots, x_i). \end{cases}$$

① Construction of v_1

We take $v_1 = \dfrac{x_1}{\|x_1\|}$: $\|v_1\| = \dfrac{\|x_1\|}{\|x_1\|} = 1$

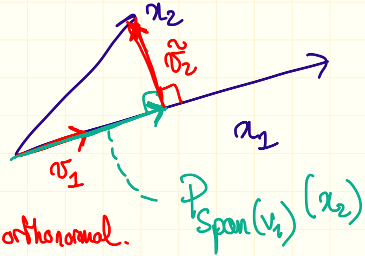
$$\left.\begin{array}{l} \bullet\ (v_1) \text{ is orthonormal} \\ \bullet\ \mathrm{Span}(v_1) = \mathrm{Span}(x_1) \end{array}\right\} \mathcal{H}_1$$

Iterative construction of the v_i's

② Construction of v_2

$$\tilde{v}_2 = x_2 - \underbrace{P_{\mathrm{Span}(v_1)}(x_2)}_{\langle v_1, x_2 \rangle v_1}$$

because (v_1) is orthonormal.



- $\langle \tilde{v}_2, v_1 \rangle = \langle v_1, x_2 \rangle - \langle v_1, x_2 \rangle \underbrace{\langle v_1, v_1 \rangle}_{=1} = 0$
- Let's define $v_2 = \dfrac{\tilde{v}_2}{\|\tilde{v}_2\|}$
- $\|v_1\| = 1$, $\|v_2\| = 1$ and $\langle v_1, v_2 \rangle = \left\langle v_1, \dfrac{\tilde{v}_2}{\|\tilde{v}_2\|} \right\rangle = 0$

(v_1, v_2) is orthonormal

Iterative construction of the v_i's

$$\mathrm{Span}(v_1, v_2) = \mathrm{Span}(x_1, x_2)$$

We constructed v_1, v_2 as linear combination of x_1 and x_2:

$$\mathrm{Span}(v_1, v_2) \subset \mathrm{Span}(x_1, x_2)$$

- $\dim \mathrm{Span}(v_1, v_2) = 2$ because (v_1, v_2) lin indep.
- $\dim \mathrm{Span}(x_1, x_2) = 2$ —— (x_1, x_2) ——

$$\boxed{\mathrm{Span}(v_1, v_2) = \mathrm{Span}(x_1, x_2)}$$

Iterative construction of the v_i's

More generally, assuming that we constructed $v_1 \dots v_i$ verifying $\mathcal{H}_i$, let's construct v_{i+1}

- Let: $$\tilde{v}_{i+1} = x_{i+1} - \overbrace{P_{\mathrm{Span}(v_1 \dots v_i)}(x_{i+1})}^{\langle v_1, x_{i+1}\rangle v_1 + \dots + \langle v_i, x_{i+1}\rangle v_i}$$

- We know that $\tilde{v}_{i+1} \perp \mathrm{Span}(v_1, \dots, v_i)$
- We let $v_{i+1} = \dfrac{\tilde{v}_{i+1}}{\|\tilde{v}_{i+1}\|}$

$(v_1 \dots v_{i+1})$ is orthonormal.

- $\mathrm{Span}(v_1 \dots v_{i+1}) = \mathrm{Span}(x_1 \dots x_{i+1})$ Exercise: check it!