MAD 4471: Introduction to Cryptography and Coding Theory **Fall 2022**

Lecture 15: The Gram-Schmidt Orthogonalization Process

Lecturer: Jean-François Biasse *TA: William Youmans*

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications. They may be distributed outside this class only with the permission of the Instructor.*

15.1 Orthogonal vectors

In the context of ideal lattices, we will consider $\mathbb{R}$-vectors spaces. Orthogonality of vectors is defined from an *inner product* over real of complex vectors spaces. In this lecture, we only state the relevant definitions for real vector spaces.

Definition 15.1 (Inner product over a real vector space) *Let V be an $\mathbb{R}$-vector space. An inner product is a map $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ satisfying*

- *Symmetry: $\forall x, y \in V^2$, $\langle x, y \rangle = \langle y, x \rangle$.*
- *Linearity: $\forall x, y, z \in V^2$, $\forall \lambda \in \mathbb{R}$, $\langle \lambda x + y, z \rangle = \lambda \langle x, y \rangle + \langle y, z \rangle$.*
- *Positive-definitness: $\forall x \neq 0$, $\langle x, x \rangle > 0$.*

Example 1 (Dot product over $\mathbb{R}^n$) *The most common example of an inner product is the dot product over $\mathbb{R}^n$. Let $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$, we have*

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n.$$

Can you think of non-trivial inner products? The definition yields a standard property which is rather straightforward for the dot product:

Proposition 15.2 *Let V be an $\mathbb{R}$-vector space and $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ be an inner product.*

- *$\langle \mathbf{0}, x \rangle = \langle x, \mathbf{0} \rangle = 0$.*
- *$\forall x \in V$, $\langle x, x \rangle = 0$ if and only if $x = 0$.*

The proof of the above statement is easy and left as an exercise. It allows us to discuss the notion of Euclidean norm.

Proposition 15.3 (Euclidean norm) *Let V be an $\mathbb{R}$-vector space and $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ be an inner product. For all $x \in V$, we define*

$$\|x\| = \sqrt{\langle x, x \rangle}.$$

This is a norm, which means that it satisfies

- $\forall x \in V$, $\|x\| = 0$ *if and only if* $x = 0$.
- $\forall x \in V, \lambda \in \mathbb{R}$, $\|\lambda x\| = |\lambda| \|x\|$.
- $\forall x, y \in V$, $\|x + y\| \leq \|x\| + \|y\|$.

Proof: The first and second properties are a direct consequence of the definition of an inner product. The third one is non-trivial to prove. It requires us to first prove the *Cauchy-Schwartz inequality*

$$\forall x, y \in V,\ \langle x, y\rangle^2 \leq \langle x, x\rangle \cdot \langle y, y\rangle.$$

Let's prove it. If $x = 0$, the statement is trivial. Let $x, y \in V$ with $x \neq 0$. We define $p : \mathbb{R} \to \mathbb{R}$ by $p(t) = \langle tx + y, tx + y\rangle$. We expand $p(t)$ by bilinearity to obtain

$$p(t) = \|x\|^2 t^2 + t(\langle x, y\rangle + \langle y, x\rangle + \|y\|^2 = \|x\|^2 t^2 + 2t\langle x, y\rangle + \|y\|^2.$$

Therefore, $p(t)$ is a degree 2 polynomial. Since we have that $p(t) \geq 0$ all the time, $p(t)$ never changes sign. Hence its discriminant Δ cannot be positive:

$$\Delta = 4\left(\langle x, y\rangle^2 - \|x\|^2 \|y\|^2\right) \leq 0.$$

This proves that $\langle x, y\rangle^2 \leq \|x\|^2 \|y\|^2$. This in turn implies the desired triangular inequality:

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y\rangle = \langle x, x\rangle + \langle y, y\rangle + 2\langle x, y\rangle \\ &= \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle \\ &\leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| \\ &= (\|x\| + \|y\|)^2 . \end{aligned}$$

Since $\|z\| \geq 0$ for all $z \in V$, we can take square roots on both sides and obtain $\|x + y\| \leq \|x\| + \|y\|$. ■

Corollary 15.4 (Euclidean distance) *The function* $d : V \times V \to \mathbb{R}_+$ *defined by* $d(x, y) = \|x - y\| = \sqrt{\langle x - y, x - y\rangle}$ *is a distance on* V*, i.e. it satisfies*

- $\forall x \in V$, $d(x, x) = 0$.
- $\forall x, y \in V$, *if* $x \neq y$, $d(x, y) > 0$.
- $\forall x, y \in V$, $d(x, y) = d(y, x)$.
- $\forall x, y, z \in V$, $d(x, z) \leq d(x, y) + d(y, z)$.

With that in mind, we can define orthogonal and orthonormal bases of V.

Definition 15.5 *We say that* $x, y \in V$ *are orthogonal if* $\langle x, y\rangle = 0$*. Moreover we define the following properties for bases of* V *(stated in the case of* V *a finite dimensional vector space):*

- *A basis* $v_1, \ldots, v_n$ *is orthogonal if* $\langle v_i, v_j\rangle = 0$ *for* $i \neq j$.
- *A basis* $v_1, \ldots, v_n$ *is orthonormal if it is orthogonal, and if* $\forall i, \|v_i\| = 1$.

15.2 Orthogonal projections

The concept of an orthogonal projection can be explained from orthonormal bases. However, we have not formally proved that they exist. For example, with the dot product of $\mathbb{R}^n$, the canonical basis $e_1 = (1, 0 \ldots, 0), \ldots, e_n = (0, \ldots, 0, 1)$ is orthonormal.

Proposition 15.6 *Let V be a finite dimensional $\mathbb{R}$-vector space. There is an orthogonal basis of V.*

Proof: We proceed by induction on the dimension n of V. For $n = 1$, a basis of V has only 1 element, and it is therefore orthogonal. Now let $n \leq 1$, and assume that all n-dimensional $\mathbb{R}$-vector space, has an orthogonal basis.
Let V be an $n+1$-dimensional $\mathbb{R}$-vector space with basis $v_1, \ldots, v_{n+1}$, and let $V' = \mathrm{Span}(v_1, \ldots, v_n)$ be the dimension-n subspace of V spanned by the first n vectors of the basis of V. From the induction hypothesis, we know that V' has an orthogonal basis $v'_1, \ldots, v'_n$. Let $x \in V \setminus V'$. We define

$$v'_{n+1} = x - \sum_i \frac{\langle x, v_i \rangle}{\langle v_i, v_i \rangle} v_i.$$

Such an element is orthogonal to all v_j for $j \leq n$. Indeed, we have for all $j \leq n$

$$\langle v'_{n+1} v_j \rangle = \langle x, v_j \rangle - \sum_i \frac{\langle x, v_i \rangle}{\langle v_i, v_i \rangle} \langle v_i, v_j \rangle = \langle x, v_j \rangle - \frac{\langle x, v_j \rangle}{\langle v_j, v_j \rangle} \langle v_j, v_j \rangle = \langle x, v_j \rangle - \langle x, v_j \rangle = 0.$$

Moreover, since $x \notin V'$, we have that $v'_{n+1} \notin V'$ as well, since otherwise, $v'_{n+1} + \sum_i \frac{\langle x, v_i \rangle}{\langle v_i, v_i \rangle} v_i = x$ would be in V'. Therefore, the dimension of $\mathrm{Span}(v'_1, \ldots, v'_{n+1})$ is $n+1$, which proves that $v'_1, \ldots, v'_{n+1}$ is a basis of V. We have proved that all $n+1$-dimensional $\mathbb{R}$-vector spaces had an orthogonal basis. By induction, all finite dimensional $\mathbb{R}$-vector spaces have an orthogonal basis. ■

Corollary 15.7 *Let V be a finite dimensional $\mathbb{R}$-vector space. There is an orthonormal basis of V.*

Proof: Let $v_1, \ldots, v_n$ be an orthogonal basis of V. For each i, we define $v'_i = \frac{v_i}{\|v_i\|}$. Then $v'_1, \ldots, v'_n$ is orthogonal, and in addition, $\|v'_i\| = 1$ for all i. ■

Definition 15.8 (Orthogonal projection) *Let V be an $\mathbb{R}$-vector space, and let $V' \subseteq V$ be a dimension-n sub-vector space of V. Let $v'_1, \ldots, v'_n$ be an orthonormal basis of V'. We define $P_{V'} : V \to V'$ by*

$$P_{V'}(x) = \sum_i \langle v'_i, x \rangle v'_i.$$

Note that the definition of $P_{V'}$ is indenpendent from the particular basis $v'_1, \ldots, v'_n$ that we choose. The following property shows that an orthogonal projection is a general projection.

Proposition 15.9 *Let V be an $\mathbb{R}$-vector space, and let $V' \subseteq V$ be a dimension-n sub-vector space of V. We have $P_{V'} \circ P_{V'}(x) = P_{V'}(x)$. Additionally, if $x \in V'$, then $P_{V'}(x) = x$.*

Proof: Let $y = P_{V'}(x)$. For each $j \leq n$, we have

$$\langle v'_j, y \rangle = \sum_i \langle v'_i, x \rangle \cdot \langle v'_i, v'_j \rangle = \langle v'_j, x \rangle \cdot \langle v'_j, v'_j \rangle = \langle v'_j, x \rangle.$$

Hence,

$$P_{V'}(y) = \sum_i \langle v'_i, y\rangle v'_i = \sum_i \langle v'_i, x\rangle v'_i = P_{V'}(x).$$

■

Example 2 *Let $V = \mathbb{R}^2$, and $V' = \mathrm{Span}((1,0)) = \{\mathbf{x} \in \mathbb{R}^2$ such that $\mathbf{x} = (x_1, 0)\}$. Then for $\mathbf{x} = (x_1, x_2)$, $P_{V'}(\mathbf{x}) = (x_1, 0)$.*

Definition 15.10 (Orthogonal complement) *Let V be an $\mathbb{R}$-vector space and $V' \subseteq V$ be a sub-vector space of V. We define the orthogonal complement of $V'^{\perp}$ of V' by*

$$V'^{\perp} = \{x \in V \mid \quad \forall v' \in V' \quad \langle x, v'\rangle = 0\}$$

We can decompose an element $x \in V$ as a sum $x = x_1 + x_2$ of an element $x_1 \in V'$ and $x_2 \in V'^{\perp}$ in a unique way. In the following proposition, we prove this fact when V is of finite dimension by using our orthogonal projections. Note that these results extend in infinite dimension, but we do not need this generalization. Our use case is Euclidean lattices, which have finite dimension.

Proposition 15.11 *Let V be a dimension-n vector space over $\mathbb{R}$ and $V' \subseteq V$ be a subspace of V of dimension $k \leq n$. We have that*

$$V = V' \oplus V'^{\perp},$$

which means that $\forall x \in V$, there are unique $x_1 \in V'$, $x_2 \in V'^{\perp}$ such that $x = x_1 + x_2$. Furthermore, we have that

- *$x_1 = P_{V'}(x)$.*
- *$x_2 = P_{V'^{\perp}}(x)$.*
- *$\dim(V'^{\perp}) = n - k$ (and thus $\dim(V') + \dim(V'^{\perp}) = \dim(V)$).*

Proof: Let $v'_1, \ldots, v'_k$ be an orthonormal basis of V', and let $x \in V$. we have that

$$x = P_{V'}(x) + (x - P_{V'}(x)) = \sum_{i \leq k} \langle v'_i, x\rangle v'_i + \left(x - \sum_{i \leq k} \langle v'_i, x\rangle v'_i\right).$$

Let $j \leq k$, we have

$$\langle \left(x - \sum_{i \leq k} \langle v'_i, x\rangle v'_i\right), v'_j\rangle = \langle x, v'_j\rangle - \langle x, v'_j\rangle = 0.$$

Hence $\forall y \in V'$, we have $\langle y, x - P_{V'}(x)\rangle = 0$. This means that $x - P_{V'}(x) \in V'^{\perp}$. In turn, $\forall x \in V$, there are $x_1 \in V'$ and $x_2 \in V'^{\perp}$ such that $x = x_1 + x_2$. However, these are not necessarily unique. Suppose we have another decomposition $x = x'_1 + x'_2$. Then we have $\underbrace{x_1 - x'_1}_{\in V'} = \underbrace{x_2 - x'_2}_{\in V'^{\perp}}$. This means that $x_1 - x'_1 \in V' \cap V'^{\perp}$. But if $y \in V' \cap V'^{\perp}$, then $\|y\| = \sqrt{\langle y, y\rangle} = 0$, which means $y = 0$. So $x_1 = x'_1$ and $x_2 = x'_2$. Let $v'_{k+1}, \ldots, v'_{k+\ell}$ be an orthonormal basis of $V'^{\perp}$. Since $x - P_{V'}(x) \in V'^{\perp}$ we know that $x - P_{V'}(x) = P_{V'^{\perp}}(x - P_{V'}(x)) = P_{V'^{\perp}}(x) - P_{V'^{\perp}}(\underbrace{P_{V'}(x)}_{\in V'})$. If $y \in V'$, $P_{V'^{\perp}}(y) = \sum_i \langle v'_{i+k}, y\rangle v'_{i+k} = 0$, hence $x - P_{V'}(x) = P_{V'^{\perp}}(x)$. This means that $x_1 = P_{V'}(x)$ and $x_2 = P_{V'^{\perp}}(x)$. Also, $v'_1, \ldots, v'_{k+\ell}$ is a basis of V, hence $\dim(V'^{\perp}) = n - k$. ■

Example 3 *Let* $V = \mathbb{R}^3$*, and* $V' = \mathrm{Span}((1,0,0),(0,0,1))$*. Then* $V'^{\perp} = \mathrm{Span}(0,1,0)$*, and*

$$\forall x = (x_1, x_2, x_3) \in V, x = \underbrace{(x_1, 0, x_3)}_{P_{V'}} + \underbrace{(0, x_2, 0)}_{P_{V'^{\perp}}}.$$

15.3 Gram-Schmidt orthogonalization

The Gram-Schmidt orthogonalization process consists in turning an arbitrary basis $v_1, \ldots, v_n$ of an n-dimensional $\mathbb{R}$-vector space V into an orthonormal basis $v'_1, \ldots, v'_n$. To do so, we use the inductive technique of the proof of Proposition 15.6: once k orthonormal vectors $v'_1, \ldots, v'_k$ are found, we add a $k+1$-th one by projecting an $x \in V$ onto $\mathrm{Span}(v'_1, \ldots, v'_k)$.

To describe this process, let's focus on the first steps. We pick $v'_1 = \frac{v_1}{\|v_1\|}$. Clearly, $\{v'_1\}$ is an orthonormal basis of the one-dimensional vector space $\mathrm{Span}(v_1)$. Now let us create v'_2 such that

- $\langle v'_1, v'_2 \rangle = 0$.
- $\|v'_2\| = 1$.
- $\mathrm{Span}(v'_1, v'_2) = \mathrm{Span}(v_1, v_2)$.

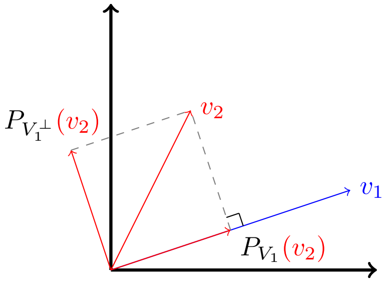

Figure 15.1: The projectors P_{V_1} and $P_{V_1^{\perp}} = I - P_{V_1}$ decompose the vector v_2 into the component $P_{V_1}(v_2)$, and the orthogonal component $P_{V_1^{\perp}}(v_2)$.

As shown in Figure 15.1, we can decompose $v_2 = P_{V_1}(v_2) + P_{V_1^{\perp}}(v_2)$, where $V_2 = \mathrm{Span}(v_1, v_2)$, and $V_1^{\perp} \subseteq V_2$ is the orthogonal complement of V_1 within V_2. We define $v'_2 = \frac{P_{V_1^{\perp}}(v_2)}{\|P_{V_1^{\perp}}(v_2)\|}$. Such an element satisfies $\|v'_2\| = 1$, and $v'_2 \in V_1^{\perp}$. On the one had, we have that $v'_1, v'_2 \in V_2$, and in addition, $v_1 = \|v_1\| v'_1 \in \mathrm{Span}(v'_1, v'_2)$, and $v_2 = \|P_{V_1}(v_2)\| v'_1 + \|P_{V_1^{\perp}}(v_2)\| v'_2 \in \mathrm{Span}(v'_1, v'_2)$. Hence $\mathrm{Span}(v'_1, v'_2) = V_2$.

Now let $V_k = \mathrm{Span}(v_1, \ldots, v_k)$, and assume that we have created an orthonormal basis $v'_1, \ldots, v'_k$ of V_k. As we did above for $k = 2$, we define

$$v'_{k+1} = \frac{P_{V_k^{\perp}}(v_{k+1})}{\|P_{V_k^{\perp}}(v_{k+1})\|} = \frac{v_{k+1} - P_{V_k}(v_{k+1})}{\|v_{k+1} - P_{V_k}(v_{k+1})\|}.$$

With that choice, we have $V_{k+1} = \mathrm{Span}(v'_1, \ldots, v'_{k+1})$, $v'_{k+1} \perp v'_i$ for $i \leq k$, $\|v'_{k+1}\| = 1$. Therefore, we have completed the next step of the process. Once $k = n$, we have an orthonormal basis of V.

Example 4 *Let $v_1 = (1,1)$ and $v_2 = (1,2)$ which are a basis of $V = \mathbb{R}^2$. The Gram-Schmid process goes like this:*

1. *Compute $v_1' = v_1/\|v_1\| = \frac{1}{\sqrt{2}}(1,1)$.*
2. *Compute $v_2' = \frac{P_{V_1^\perp}(v_2)}{\|P_{V_1^\perp}(v_2)\|}$:*
    - *Calculate $\langle v_2, v_1' \rangle = 3/\sqrt{2}$.*
    - *Calculate $P_{V_1}(v_2) = \langle v_2, v_1' \rangle v_1' = \frac{3}{2}(1,1)$ and $P_{V_1^\perp}(v_2) = v_2 - P_{V_1}(v_2) = (1,2) - \frac{3}{2}(1,1) = \frac{1}{2}(-1,1)$.*
    - *Deduce $v_2' = \frac{P_{V_1^\perp}(v_2)}{\|P_{V_1^\perp}(v_2)\|} = \frac{\frac{1}{2}(-1,1)}{\sqrt{\frac{1}{4}+\frac{1}{4}}} = \frac{1}{\sqrt{2}}(-1,1)$.*

Hence the Gram-Schmidt orthonormalization of $v_1 = (1,1), v_2 = (1,2)$ is $v_1' = \frac{1}{\sqrt{2}}(1,1), v_2' = \frac{1}{\sqrt{2}}(-1,1)$. Note that this is not the most straightforward orthonormal basis of $V = \mathbb{R}^2$. Indeed, one could pick $(1,0),(0,1)$. However the Gram-Schmidt process is only guaranteed to return one of the infinitely many bases of V.