Lecture Notes: The Gram-Schmidt Orthogonalization Process

Linear Algebra - MA327

Heitor Baldo*

May 5th, 2016

1 Preliminary Concepts

Throughout this text, V denotes a vector space over a field $\mathbb{F}$, where $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$. The primary bibliography is [1, 2, 3, 4].

Definition 1.1. An *inner product* on V is a map $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$ satisfying, for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and scalars $\alpha, \beta \in \mathbb{F}$:

1. $\langle \alpha\mathbf{u} + \beta\mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{u}, \mathbf{w} \rangle + \beta \langle \mathbf{v}, \mathbf{w} \rangle$;
2. $\langle \mathbf{u}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{u} \rangle}$;
3. $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$, with $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ iff $\mathbf{v} = \mathbf{0}$.

A vector space V equipped with an inner product $\langle \cdot, \cdot \rangle$ is called an *inner product space*, and we write $(V, \langle \cdot, \cdot \rangle)$.

The inner product induces a *norm* $\|\mathbf{v}\| := \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$, which measures length in the space.

Example 1.2. Consider the n-dimensional Euclidean space $\mathbb{R}^n$; for $\mathbf{u} = (u_1, ..., u_n)$, $\mathbf{v} = (v_1, ..., v_n) \in \mathbb{R}^n$, the map $\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{i=1}^n u_i v_i \in \mathbb{R}$ define an inner product in $\mathbb{R}^n$ (this inner product is also called *dot product*). Consider now the space $C[a, b]$ of continuous functions; for $f, g \in C[a, b]$, the map $\langle f, g \rangle = \int_a^b f(x)\, g(x)\, dx \in \mathbb{R}$ defines an inner product in this space.

Definition 1.3. Given two vectors $\mathbf{u}, \mathbf{v} \in V$, they are called *orthogonal* if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$, and we write $\mathbf{u} \perp \mathbf{v}$. A set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is called

- *orthogonal* if $\left\langle \mathbf{v}_i, \mathbf{v}_j \right\rangle = 0$, for all $i, j = 1, .., n$, $i \neq j$;

*hbaldo@ime.unicamp.br

- *orthonormal* if it is orthogonal and $\|\mathbf{v}_i\| = 1$ for all $i = 1, .., n$, that is $\left\langle \mathbf{v}_i, \mathbf{v}_j \right\rangle = \delta_{ij}$.

Proposition 1.4. *Any orthogonal set of nonzero vectors is linearly independent.*

Proof. Suppose $\mathcal{U} = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is an orthogonal set of nonzero vectors ($\mathbf{u}_i \neq \mathbf{0}$, $\mathbf{u}_i \in V$) and suppose $\sum_{i=1}^{k} c_i \mathbf{u}_i = \mathbf{0}$, $c_i \in \mathbb{F}$. For an arbitrary $j \in \{1, \ldots, k\}$, we have:

$$0 = \left\langle \sum_i c_i \mathbf{u}_i, \mathbf{u}_j \right\rangle = c_j \left\langle \mathbf{u}_j, \mathbf{u}_j \right\rangle = c_j \left\|\mathbf{u}_j\right\|^2 .$$

Since $\mathbf{u}_j \neq \mathbf{0}$ we have $\left\|\mathbf{u}_j\right\|^2 > 0$, forcing $c_j = 0$. As j was arbitrary, all coefficients c_j vanish. Therefore, $\mathcal{U}$ is linearly independent. □

2 Orthogonal Projections

The idea behind orthogonal projection is that, given a nonzero vector $\mathbf{u}$, we want to decompose any $\mathbf{v}$ into a part parallel to $\mathbf{u}$ and a part perpendicular to it.

Definition 2.1. Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space. Given vectors $\mathbf{u}, \mathbf{v} \in V$, $\mathbf{u} \neq \mathbf{0}$, the *orthogonal projection* of $\mathbf{v}$ onto the line spanned by $\mathbf{u}$ is defined by

$$\operatorname{proj}_{\mathbf{u}}(\mathbf{v}) = \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \mathbf{u}.$$

Lemma 2.2. *For any $\boldsymbol{v}$ and nonzero $\boldsymbol{u}$, the residual $\boldsymbol{r} = \boldsymbol{v} - \operatorname{proj}_{\boldsymbol{u}}(\boldsymbol{v})$ satisfies $\boldsymbol{r} \perp \boldsymbol{u}$.*

Proof. Computing the inner product $\langle \mathbf{r}, \mathbf{u} \rangle$, we have:

$$\langle \mathbf{r}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle - \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \langle \mathbf{u}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle - \langle \mathbf{v}, \mathbf{u} \rangle = 0.$$ □

More generally, if $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ is an orthogonal set spanning a subspace $W \subseteq V$, the projection of $\mathbf{v}$ onto W is the sum of the one-dimensional projections, that is,

$$\operatorname{proj}_W(\mathbf{v}) = \sum_{j=1}^{m} \frac{\left\langle \mathbf{v}, \mathbf{u}_j \right\rangle}{\left\langle \mathbf{u}_j, \mathbf{u}_j \right\rangle} \mathbf{u}_j,$$

and $\mathbf{v} - \operatorname{proj}_W(\mathbf{v})$ is orthogonal to *all* of W. This allows the Gram-Schmidt process to build an orthogonal set incrementally as we will see in the next section.

3 The Gram-Schmidt Process

Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space and let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be linearly independent vectors in V.

Define orthogonal vectors $\mathbf{u}_1, \dots, \mathbf{u}_n$ recursively:

$$\mathbf{u}_1 = \mathbf{v}_1, \tag{1}$$

$$\mathbf{u}_k = \mathbf{v}_k - \sum_{j=1}^{k-1} \frac{\langle \mathbf{v}_k, \mathbf{u}_j \rangle}{\langle \mathbf{u}_j, \mathbf{u}_j \rangle} \mathbf{u}_j, \qquad k = 2, 3, \dots, n. \tag{2}$$

Then normalize to obtain an orthonormal set:

$$\mathbf{e}_k = \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|}, \qquad k = 1, \dots, n.$$

The vectors $\mathbf{u}_1, \dots, \mathbf{u}_n$ are all nonzero, pairwise orthogonal, and satisfy the following property

$$\operatorname{span}\{\mathbf{u}_1, \dots, \mathbf{u}_k\} = \operatorname{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\} \qquad \text{for every } k = 1, \dots, n.$$

Consequently $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is an orthonormal basis of $\operatorname{span}\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$.

In fact, by induction on k, for $k = 1$ we have $\mathbf{u}_1 = \mathbf{v}_1 \neq \mathbf{0}$ (as $\mathbf{v}_1$ belongs to an independent set), and trivially $\operatorname{span}\{\mathbf{u}_1\} = \operatorname{span}\{\mathbf{v}_1\}$. Now, assume that the claim holds for $k-1$: the vectors $\mathbf{u}_1, \dots, \mathbf{u}_{k-1}$ are nonzero, pairwise orthogonal, and span the same space as $\mathbf{v}_1, \dots, \mathbf{v}_{k-1}$. Fix any index $i < k$ and take the inner product of the formula (2) for $\mathbf{u}_k$ with $\mathbf{u}_i$. Because the $\mathbf{u}_j$ ($j < k$) are mutually orthogonal, every term in the sum (2) vanishes except $j = i$, and then we have:

$$\langle \mathbf{u}_k, \mathbf{u}_i \rangle = \langle \mathbf{v}_k, \mathbf{u}_i \rangle - \frac{\langle \mathbf{v}_k, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \langle \mathbf{u}_i, \mathbf{u}_i \rangle = 0.$$

Thus $\mathbf{u}_k$ is orthogonal to each $\mathbf{u}_i$, extending orthogonality to the set $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$. Next, $\mathbf{u}_k \neq \mathbf{0}$. Were it zero, the formula (2) would express $\mathbf{v}_k$ as a linear combination of $\mathbf{u}_1, \dots, \mathbf{u}_{k-1}$, which by the inductive hypothesis lie in $\operatorname{span}\{\mathbf{v}_1, \dots, \mathbf{v}_{k-1}\}$; that would make $\mathbf{v}_k$ dependent on its predecessors, contradicting the independence of the set.

Finally, the formula (2) writes $\mathbf{u}_k$ as $\mathbf{v}_k$ plus a combination of $\mathbf{u}_1, \dots, \mathbf{u}_{k-1}$, hence (by induction) plus a combination of $\mathbf{v}_1, \dots, \mathbf{v}_{k-1}$; therefore $\mathbf{u}_k \in \operatorname{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$, giving the inclusion $\operatorname{span}\{\mathbf{u}_1, \dots, \mathbf{u}_k\} \subseteq \operatorname{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$. Both sides have dimension k (the left because the orthogonal $\mathbf{u}_i$ are independent by Proposition 1.4, and the right by hypothesis, so the inclusion is actually an equality). This completes the induction. Normalizing preserves orthogonality and span while rescaling each vector to unit length, so $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is the desired orthonormal basis.

4 Examples

Example 4.1. In this example, let's orthogonalize three vectors in $\mathbb{R}^3$ equipped with the standard dot product (Example 1.2). Let's apply the Gram-Schmidt process in the following vectors:

$$\mathbf{v}_1 = \begin{pmatrix}1\\1\\0\end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix}1\\0\\1\end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix}0\\1\\1\end{pmatrix}.$$

Step 1. Take $\mathbf{u}_1 = \mathbf{v}_1 = (1,1,0)^\top$, with $\langle \mathbf{u}_1,\, \mathbf{u}_1\rangle = 2$.

Step 2. Since $\langle \mathbf{v}_2,\, \mathbf{u}_1\rangle = 1\cdot 1 + 0\cdot 1 + 1\cdot 0 = 1$,

$$\mathbf{u}_2 = \mathbf{v}_2 - \frac{1}{2}\,\mathbf{u}_1 = \begin{pmatrix}1\\0\\1\end{pmatrix} - \frac{1}{2}\begin{pmatrix}1\\1\\0\end{pmatrix} = \begin{pmatrix}\frac{1}{2}\\-\frac{1}{2}\\1\end{pmatrix}, \qquad \langle \mathbf{u}_2,\, \mathbf{u}_2\rangle = \tfrac{1}{4} + \tfrac{1}{4} + 1 = \tfrac{3}{2}.$$

Step 3. We need $\langle \mathbf{v}_3,\, \mathbf{u}_1\rangle = 0\cdot 1 + 1\cdot 1 + 1\cdot 0 = 1$ and $\langle \mathbf{v}_3,\, \mathbf{u}_2\rangle = 0\cdot\frac{1}{2} + 1\cdot(-\frac{1}{2}) + 1\cdot 1 = \frac{1}{2}$. Hence

$$\mathbf{u}_3 = \mathbf{v}_3 - \frac{1}{2}\,\mathbf{u}_1 - \frac{1/2}{3/2}\,\mathbf{u}_2 = \begin{pmatrix}0\\1\\1\end{pmatrix} - \frac{1}{2}\begin{pmatrix}1\\1\\0\end{pmatrix} - \frac{1}{3}\begin{pmatrix}\frac{1}{2}\\-\frac{1}{2}\\1\end{pmatrix} = \begin{pmatrix}-\frac{2}{3}\\\frac{2}{3}\\\frac{2}{3}\end{pmatrix}.$$

One verifies $\langle \mathbf{u}_1,\, \mathbf{u}_3\rangle = -\frac{2}{3} + \frac{2}{3} + 0 = 0$ and $\langle \mathbf{u}_2,\, \mathbf{u}_3\rangle = -\frac{1}{3} - \frac{1}{3} + \frac{2}{3} = 0$, as expected.

Normalize. With $\|\mathbf{u}_1\| = \sqrt{2}$, $\|\mathbf{u}_2\| = \sqrt{3/2} = \frac{\sqrt{6}}{2}$, and $\|\mathbf{u}_3\| = \sqrt{4/3} = \frac{2}{\sqrt{3}}$,

$$\mathbf{e}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\end{pmatrix}, \qquad \mathbf{e}_2 = \frac{1}{\sqrt{6}}\begin{pmatrix}1\\-1\\2\end{pmatrix}, \qquad \mathbf{e}_3 = \frac{1}{\sqrt{3}}\begin{pmatrix}-1\\1\\1\end{pmatrix}.$$

These three unit vectors are mutually perpendicular and span all of $\mathbb{R}^3$ (orthonormal basis).

Example 4.2. Let V be the space of polynomials on $[-1,1]$ with the inner product $\langle f,\, g\rangle = \int_{-1}^{1} f(x)\,g(x)\,dx$. Let's apply the Gram-Schmidt process to orthogonalize the monomial basis $v_0 = 1,\ v_1 = x,\ v_2 = x^2$.

Step 0. $u_0 = 1$, with $\langle u_0,\, u_0\rangle = \int_{-1}^{1} 1\,dx = 2$.

Step 1. Since $\langle x,\, 1\rangle = \int_{-1}^{1} x\,dx = 0$ (an odd integrand over a symmetric interval), the projection vanishes and

$$u_1 = x - \frac{0}{2}\cdot 1 = x, \qquad \langle u_1,\, u_1\rangle = \int_{-1}^{1} x^2\,dx = \tfrac{2}{3}.$$

Step 2. We need $\langle x^2, 1\rangle = \int_{-1}^{1} x^2\,dx = \frac{2}{3}$ and $\langle x^2, x\rangle = \int_{-1}^{1} x^3\,dx = 0.$ Therefore

$$u_2 = x^2 - \frac{2/3}{2}\cdot 1 - \frac{0}{2/3}\cdot x = x^2 - \frac{1}{3}.$$

The polynomials $1,\ x,\ x^2 - \frac{1}{3}$ are mutually orthogonal under this integral inner product.

References

[1] Halmos, P. (1950). Finite-Dimensional Vector Spaces, Springer Verlag.

[2] Hoffman, K. M., & Kunze, R. (1971). Linear Algebra. 2nd ed., Pearson.

[3] Kostrikin, A. I., & Manin, Y. I. (1997). Linear Algebra. CRC Press.

[4] Lang, S. (1987). Linear Algebra, Springer Verlag.