# Online Lecture Notes

Prof. Boris Houska

March 15, 2022

## 1 Gram-Schmidt Algorithm

The goal of this section is to apply the Gram-Schmidt algorithm to the monomial basis functions

$$a_0(x) = 1,\ a_1(x) = x,\ a_2(x) = x^2,\ \ldots$$

in the Hilbert space $L_2[-1,1]$. Here, $L_2[-1,1]$ denotes the set of square-integrable functions on the interval $[-1,1]$ with respect to the scalar product

$$\langle f, g \rangle \ = \ \int_{-1}^{1} f(x)g(x)\,\mathrm{d}x$$

and the corresponding norm

$$\|f\|_H = \sqrt{\int_{-1}^{1} f(x)^2\,\mathrm{d}x}\ .$$

This is in analogy to the standard scalar product and Euclidean norm in $\mathbb{R}^n$, but after replacing sums with integrals.

### 1.1 First iteration of the Gram-Schmidt Algorithm

We need to start with the vector

$$\bar{q}_0(x) = a_0(x) = 1\ .$$

We can proceed with the normalization step

$$q_0(x) = \frac{\bar{q}_0(x)}{\|\bar{q}_0\|_H} = \frac{1}{\sqrt{\int_{-1}^{1} 1\,\mathrm{d}x}} = \frac{1}{\sqrt{2}}\ .$$

Notice that this construction is such that

$$\|q_0\|_H = 1\ .$$

## 1.2 Second iteration of the Gram-Schmidt Algorithm

We are back to the next orthogonalization step:

$$\bar{q}_1(x) = a_1(x) - \langle a_1, q_0 \rangle q_0(x) = x - \underbrace{\left[ \int_{-1}^{1} x \frac{1}{\sqrt{2}} \, \mathrm{d}x \right]}_{=0} q_0(x) = x$$

Next, we need to do yet another normalization step,

$$q_1(x) = \frac{\bar{q}_1(x)}{\|\bar{q}_1\|_H} = \frac{x}{\sqrt{\int_{-1}^{1} x^2 \, \mathrm{d}x}} = \frac{x}{\sqrt{\frac{2}{3}}} = \sqrt{\frac{3}{2}} x$$

## 1.3 Third iteration of the Gram-Schmidt Algorithm

Ok, this is getting boring, but the third orthogonalization step is given by

$$\begin{aligned}
\bar{q}_2(x) &= a_2(x) - \langle a_2, q_0 \rangle q_0(x) - \langle a_2, q_1 \rangle q_1(x) \\
&= x^2 - \underbrace{\left[ \int_{-1}^{1} x^2 \frac{1}{\sqrt{2}} \, \mathrm{d}x \right] \frac{1}{\sqrt{2}}}_{=\frac{1}{3}} - \underbrace{\left[ \int_{-1}^{1} x^2 \sqrt{\frac{3}{2}} x \, \mathrm{d}x \right]}_{=0} \sqrt{\frac{3}{2}} x \\
&= x^2 - \frac{1}{3} \qquad (1)
\end{aligned}$$

Finally, we need to do another normalization step (even more boring...)

$$q_2(x) = \frac{\bar{q}_2(x)}{\|\bar{q}_2\|_H} = \frac{x^2 - \frac{1}{3}}{\sqrt{\int_{-1}^{1} \left[ x^2 - \frac{1}{3} \right]^2 \, \mathrm{d}x}} = \sqrt{\frac{5}{8}} (3x^2 - 1) \ .$$

We can keep on doing this for all monomials. The integrals get more complicated, but this is straightforward. At the end of the day, this yields a complete orthonormal basis, $q_0, q_1, q_2$, and so on, which satisfies:

1. *Orthogonality.* We have $\langle q_i, q_j \rangle = 0$ if $i \neq j$.
2. *Normality.* We have $\sqrt{\langle q_i, q_i \rangle} = \|q_i\|_H = 1$.
3. *Completeness.* There exists for every polynomial

   $$p = \lambda_0 a_0 + \lambda_1 a_1 + \lambda_2 a_2 + \ldots + \lambda_n a_n$$

   coefficients $\kappa_0, \kappa_1, \ldots, \kappa_n \in \mathbb{R}$ such that

   $$p = \kappa_0 q_0 + \kappa_1 q_1 + \kappa_2 q_2 + \ldots + \kappa_n q_n$$

The functions $q_0, q_1, q_2 \ldots$ are called the Legendre polynomials. They are orthonormal (see above).

## 1.4 Consequences of the above construction:

If $p(x) = \lambda_0 + \lambda x + \lambda_2 x^2 + \ldots + \lambda_n x^n$ is a polynomial of order $n$, then we have

$$\begin{aligned}
\langle p, q_k \rangle &= \langle \kappa_0 q_0 + \kappa_1 q_1 + \kappa_2 q_2 + \ldots + \kappa_n q_n, q_k \rangle & (2) \\
&= \sum_{i=0}^{n} \kappa_i \langle q_i, q_k \rangle = \sum_{i=0}^{n} \kappa_i \delta_{i,k} = \begin{cases} \kappa_k & \text{if } k \leq n \\ 0 & \text{otherwise} \end{cases} , & (3)
\end{aligned}$$

for $k \in \mathbb{N}$, where $\delta_{i,k} = \begin{cases} 1 & \text{if } i = k \\ 0 & \text{otherwise} \end{cases}$ .

The above property is the basis for Gauss approximation, which we will discuss in the following lecture.

## 1.5 Generalizations

The above procedure can be implemented for any inner product in the function space $L_2[a, b]$. In the most general case, we could have $a < b$ and $a \in \mathbb{R} \cup \{-\infty\}$ and $a \in \mathbb{R} \cup \{\infty\}$. For example, we could consider weighted scalar products of the form

$$\langle f, g \rangle = \int_a^b f(x) g(x) \omega(x) \, \mathrm{d}x$$

with a given weighting function $\omega : [a, b] \to \mathbb{R}_{++}$. Examples for this are:

1. If we set $a = -1$, $b = 1$, and $\omega(x) = 1$, then the Gram-Schmidt algorithm returns the Legendre polynomials

   $$q_0, q_1, q_2, \ldots$$

   (see above)

2. If we set $a = -\infty$, $b = \infty$, and $\omega(x) = e^{-x^2/2}$, then the Gram-Schmidt algorithm returns the Hermite polynomials. (Details will be part of our homework exercises)

3. If we set $a = -1$, $b = 1$, and $\omega(x) = \frac{1}{\sqrt{1-x^2}}$, then the Gram-Schmidt algorithm returns the Chebyshev polynomials of the first kind. (Our TA, Xvting Gao, is an expert on that... - we will have exercises on this, too)

4. ... So, basically, whenever we take a different weighting factor, we get a different orthogonal basis with respect to the corresponding inner product.

## 2 Proof of Gauss' optimality conditions

We consider the general functional minimization problem

$$\min_{p \in P_n} \ \|f-p\|_H^2$$

over the set of polynomials $P_n$ of order $n$ (or smaller). The key idea is to analyze the auxiliary function

$$F(t) = \|f-p-tq\|_H^2$$

for a perturbation direction $q \in P_n$ with scalar parameter $t \in \mathbb{R}$. (This idea is borrowed from variational analysis). Here, we can work with any Hilbert space! Main observation: if $p$ is a minimizer, then $t=0$ must be a minimizer of $F$, which implies that

$$\begin{aligned}
0 &= \frac{\mathrm{d}F(0)}{\mathrm{d}t} \ = \ \frac{\mathrm{d}}{\mathrm{d}t}\|f-p-tq\|_H^2\Big|_{t=0} \ = \ \frac{\mathrm{d}}{\mathrm{d}t}\langle f-p-tq, f-p-tq\rangle\Big|_{t=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\left[\langle f-p,f-p\rangle - 2t\langle f-p,q\rangle + t^2\langle q,q\rangle\right]_{t=0} \\
&= -2\langle f-p,q\rangle \qquad (4)
\end{aligned}$$

This means that I can divide by $-2$ in order to get the optimality condition

$$\forall q \in P_n, \qquad \langle f-p,q\rangle = 0 \ .$$

The other way around, we may assume that $p$ satisfies $\langle f-p,q\rangle = 0$ for all $q \in P_n$. Then we have

$$\begin{aligned}
\|f-p\|_H^2 &= \langle f-p,f-p\rangle = \langle f-p,f-q+q-p\rangle \\
&= \langle f-p,f-q\rangle + \underbrace{\langle f-p,q-p\rangle}_{=0} \\
&= \langle f-p,f-q\rangle. \qquad (5)
\end{aligned}$$

The last step uses that $p-q$ is orthogonal to $f-p$ (since $p-q$ is a polynomial of order $\leq n$). Next, we apply the Cauchy-Schwartz inequality to find

$$\|f-p\|_H^2 \ = \ \langle f-p,f-q\rangle \ \leq \ \|f-p\|_H\|f-q\|_H \ .$$

Now, there are two cases: Case 1: $\|f-p\| = 0$. In this case $f=p$ is a polynomial and we have found the optimal solution. Case 2: we have $\|f-p\| \neq 0$. In this case, we can divide by $\|f-p\| > 0$ such that we find

$$\forall q \in P_n, \qquad \|f-p\|_H \leq \|f-q\|_H \ .$$

But this is the same as saying that $p$ is optimal!