THEORETICAL FOUNDATIONS OF GAUSSIAN QUADRATURE

1 Inner product vector space

Definition 1. *A* vector space *(or* linear space*) over $\mathbb{R}$ is a set $V = \{\mathbf{u}, \mathbf{v}, \mathbf{w}, \ldots\}$ in which the following two operations are defined:*

(A) *Addition of vectors:*
$$+ : V \times V \to V : (\mathbf{u}, \mathbf{v}) \mapsto \mathbf{u} + \mathbf{v} \ ,$$

    which satisfies the properties

    (A_1) *associativity:* $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w} \quad \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V$*;*

    (A_2) *existence of a zero vector:* $\exists \mathbf{0} \in V$ *such that* $\mathbf{u} + \mathbf{0} = \mathbf{u} \quad \forall \mathbf{u} \in V$*;*

    (A_3) *existence of an opposite element:* $\forall \mathbf{u} \in V \ \exists (-\mathbf{u}) \in V$ *such that* $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$*;*

    (A_4) *commutativity:* $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u} \quad \forall \mathbf{u}, \mathbf{v} \in V$*;*

(B) *Multiplication of a number and a vector:*
$$\mathbb{R} \times V \to V : (\alpha, \mathbf{u}) \mapsto \alpha \mathbf{u} \ ,$$

    which satisfies the properties

    (M_1) $(\alpha\beta)\mathbf{u} = \alpha(\beta\mathbf{u}) \quad \forall \alpha, \beta \in \mathbb{R}, \ \forall \mathbf{u} \in V$*;*

    (M_2) *distributivity:* $(\alpha + \beta)\mathbf{u} = \alpha\mathbf{u} + \beta\mathbf{u} \quad \forall \alpha, \beta \in \mathbb{R}, \ \forall \mathbf{u} \in V$*;*

    (M_3) *distributivity:* $\alpha(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v} \quad \forall \alpha \in \mathbb{R}, \ \forall \mathbf{u}, \mathbf{v} \in V$*;*

    (M_4) *normalization:* $1\mathbf{u} = \mathbf{u} \quad \forall \mathbf{u} \in V$*.*

Definition 2. *A* normed linear space *is a linear space V with an operation*
$$\| \ \| : V \to \mathbb{R} : \mathbf{u} \mapsto \|\mathbf{u}\| \ ,$$

satisfying the properties

(N_1) $\|\alpha\mathbf{u}\| = |\alpha| \|\mathbf{u}\| \quad \forall \alpha \in \mathbb{R}, \ \forall \mathbf{v} \in V$*;*

(N_2) *triangle inequality:* $\|\mathbf{u} + \mathbf{v}\| \le \|\mathbf{u}\| + \|\mathbf{v}\| \quad \forall \mathbf{u}, \mathbf{v} \in V$*;*

(N_3) $\|\mathbf{u}\| \ge 0 \ \forall \mathbf{u} \in V$*; moreover, if $\|\mathbf{u}\| = 0$, then $\mathbf{u} = \mathbf{0}$ (nondegeneracy).*

Definition 3. *An* inner product linear space *(or* scalar product linear space*) is a linear space V with an operation*

$$\langle \ , \ \rangle : V \times V \to \mathbb{R} : (\mathbf{u}, \mathbf{v}) \mapsto \langle \mathbf{u}, \mathbf{v} \rangle \ ,$$

satisfying the properties

(I_1) *symmetry:* $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle \quad \forall \mathbf{u} \in V$;

(I_2) *linearity:* $\langle \alpha \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle \quad \forall \alpha \in \mathbb{R}, \ \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V$;

(I_3) $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0 \quad \forall \mathbf{u} \in V$; *moreover, if* $\langle \mathbf{u}, \mathbf{u} \rangle = 0$, *then* $\mathbf{u} = \mathbf{0}$ *(nondegeneracy).*

Remark. In an inner product linear space, one can define the *norm* of a vector by

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

If we write the norm of a vector in an inner product vector space, it is <u>always</u> understood that it is defined by this equality.

Example. The "standard" or "Euclidean" inner product of the vectors $\mathbf{u} = (u_1, u_2, \ldots, u_d) \in \mathbb{R}^d$ and $\mathbf{v} = (v_1, v_2, \ldots, v_d) \in \mathbb{R}^d$ is given by

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{i=1}^{d} u_i v_i \ ,$$

and the corresponding norm is

$$\|\mathbf{u}\| = \sqrt{\sum_{i=1}^{d} u_i^2} \ .$$

Example. Let $\mathbf{G}$ be a symmetric positive-definite $d \times d$ matrix, for example

$$\mathbf{G} = (g_{ij}) = \begin{pmatrix} 5 & 4 & 1 \\ 4 & 7 & 0 \\ 1 & 0 & 3 \end{pmatrix} \ .$$

Then one can define inner product corresponding to $\mathbf{G}$ by

$$\langle \mathbf{u}, \mathbf{v} \rangle := \sum_{i=1}^{d} \sum_{j=1}^{d} u_i \, g_{ij} \, v_j \ .$$

Theorem. In any inner product linear space V, the *Cauchy-Schwarz inequality* holds:

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \, \|\mathbf{v}\| \qquad \forall \mathbf{u}, \mathbf{v} \in V \ .$$

Think about the meaning of this inequality in $\mathbf{R}^3$.

Remark. In an inner product linear space, one can define the *angle* θ between two vectors, $\mathbf{u}$ and $\mathbf{v}$, by

$$\langle \mathbf{u}, \mathbf{v} \rangle =: \|\mathbf{u}\|\|\mathbf{v}\| \cos\theta \ .$$

Note that this definition makes sense thanks to the Cauchy-Schwarz inequality.

Exercise. Find the norm of the vector $\mathbf{u} = (3, 0, -4)$ using the "standard" inner product in $\mathbf{R}^3$ and then by using the inner product in $\mathbf{R}^3$ defined through the matrix $\mathbf{G}$.

A very important example. Consider the set of all polynomials of degree no greater than 4, where the operations "addition of vectors" and "multiplication of a number and a vector" are defined in the standard way, namely: if P and Q are such polynomials,

$$P(x) = p_4 x^4 + p_3 x^3 + p_2 x^2 + p_1 x + p_0 \ , \qquad Q(x) = q_4 x^4 + q_3 x^3 + q_2 x^2 + q_1 x + q_0 \ ,$$

then their sum, $P + Q$ is given by

$$(P + Q)(x) = (p_4 + q_4)x^4 + (p_3 + q_3)x^3 + (p_2 + q_2)x^2 + (p_1 + q_1)x + (p_0 + q_0) \ ,$$

and, for $\alpha \in \mathbb{R}$, the product αP is defined by

$$(\alpha P)(x) = (\alpha p_4)x^4 + (\alpha p_3)x^3 + (\alpha p_2)x^2 + (\alpha p_1)x + (\alpha p_0) \ .$$

Then this set of polynomials is a vector space of dimension 5. One can take for a basis in this space the set of polynomials

$$E_0(x) := 1 \ , \quad E_1(x) := x \ , \quad E_2(x) := x^2 \ , \quad E_3(x) := x^3 \ , \quad E_4(x) := x^4 \ .$$

This, however, is only one of the infinitely many bases in this space. For example, the set of vectors

$$G_0(x) := x - 1 \ , \quad G_1(x) := x + 1 \ , \quad G_2(x) := x^2 + 3x + 3 \ ,$$

$$G_3(x) := -x^3 + 3x^2 - 4 \ , \quad G_4(x) := x^4 - x^3 - 2x$$

is a perfectly good basis. (Note that I called G_0, G_1, ..., G_n "vectors" to emphasize that what is important for us is the structure of vector space and not so much the fact that these "vectors" are polynomials.) Any vector (i.e., polynomial of degree ≤ 4) can be represented in a unique way in any basis, for example, the polynomial $P(x) = 3x^4 - 5x^2 + x + 7$ can be written as

$$P = 3\,E_4 - 5\,E_2 + E_1 + 7\,E_0 \ ,$$

or, alternatively, as

$$P = 3\,G_4 - 3\,G_3 + 4\,G_2 - 11\,G_1 + 6\,G_0 \ .$$

2   Inner product in the space of polynomials

One can define an inner product structure in the space of polynomials in many different ways. Let $V_n(a,b)$ stand for the space of polynomials of degree $\leq n$ defined for $x \in [a,b]$. Most of the theory we will develop works also if $a = -\infty$ and/or $b = \infty$. Let $w : [a,b] \to \mathbb{R}$ be a *weight function*, i.e., a function satisfying the following properties:

(a) the integral $\int_a^b w(x)\,\mathrm{d}x$ exists;

(b) $w(x) \geq 0$ for all $x \in [a,b]$, and $w(x)$ can be zero only at isolated points in $[a,b]$ (in particular, $w(x)$ cannot be zero in an interval of nonzero length).

We define inner product in $V_n(a,b)$ by

$$\langle P, Q \rangle := \int_a^b P(x)\,Q(x)\,w(x)\,\mathrm{d}x \; ; \tag{1}$$

if the interval (a,b) is of infinite length, then one has to take w such that this integral exists for all P and Q in $V_n(a,b)$. Let $V_n(a,b;w)$ stands for the inner product linear space of polynomials of degree $\leq n$ defined on $[a,b]$, and inner product defined by (1).

Example. The *Legendre polynomials* are a family of polynomials P_0, P_1, P_2, ... such that P_n is a polynomial of degree n defined for $x \in [-1,1]$, with leading coefficients equal to 1 ("leading" are the coefficients of the highest powers of x) and such that P_n and P_m are orthogonal for $n \neq m$ in the sense of the following inner product:

$$\langle P_n, P_m \rangle = \int_{-1}^1 P_n(x)\,P_m(x)\,\mathrm{d}x \; .$$

In other words, the polynomials P_0, P_1, P_2, ..., P_n constitute an orthogonal basis of the space $V_n(-1,1;w(x) \equiv 1)$. Here are the first several Legendre polynomials:

$$P_0(x) = 1\,, \quad P_1(x) = x\,, \quad P_2(x) = x^2 - \frac{1}{3}\,, \quad P_3(x) = x^3 - \frac{3}{5}x\,, \quad P_4(x) = x^4 - \frac{6}{7}x^2 + \frac{3}{35}\,,\dots\;.$$

Sometimes Legendre polynomials are normalized in a different way:

$$\tilde{P}_0(x) = 1\,, \quad \tilde{P}_1(x) = x\,, \quad \tilde{P}_2(x) = \frac{1}{2}(3x^2 - 1)\,,$$

$$\tilde{P}_3(x) = \frac{1}{2}(5x^3 - 3x)\,, \quad \tilde{P}_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3)\,,\dots\;;$$

check that P_n is proportional to $\tilde{P}_n$ for all the polynomials given here.

Exercise. Check that each of the first five Legendre polynomials is orthogonal to all other Legendre polynomials in the example above.

4

Example. The *Hermite polynomials* are a family of polynomials H_0, H_1, H_2, ... such that H_n is a polynomial of degree n defined for $x \in \mathbb{R}$, normalized in such a way that $\langle H_n, H_n \rangle = 2^n n! \sqrt{\pi}$ and $\langle H_n, H_m \rangle = 0$ for $n \neq m$, where the inner product is defined as follows:

$$\langle H_n, H_m \rangle = \int_{-\infty}^{\infty} H_n(x) \, H_m(x) \, \mathrm{e}^{-x^2} \, \mathrm{d}x \ .$$

In other words, the polynomials H_0, H_1, H_2, ..., H_n constitute an orthogonal basis of the space $V_n(-\infty, \infty; \mathrm{e}^{-x^2})$. Here are the first five Hermite polynomials:

$$H_0(x) = 1 \ , \ \ H_1(x) = 2x \ , \ \ H_2(x) = 4x^2 - 2 \ , \ \ H_3(x) = 8x^3 - 12x \ , \ \ H_4(x) = 16x^4 - 48x^2 + 12 \ .$$

3   Gaussian quadrature

Theorem 1. *Let w be a weight function on $[a, b]$, let n be a positive integer, and let G_0, G_1, ..., G_n be an orthogonal family of polynomials with degree of G_k equal to k for each $k = 0, 1, \ldots, n$. In other words, G_0, G_1, ..., G_n form an orthogonal basis of the inner product linear space $V_n(a, b; w)$. Let x_1, x_2, ..., x_n be the roots of G_n, and define*

$$L_i(x) := \prod_{j=1, \ j \neq i}^{n} \frac{x - x_j}{x_i - x_j} \qquad \text{for} \ \ i = 1, 2, \ldots, n \ .$$

Then the corresponding Gaussian quadrature formula is given by

$$I(f) := \int_a^b f(x) \, w(x) \, \mathrm{d}x \ \approx \ I_n(f) := \sum_{i=1}^{n} w_i \, f(x_i) \ ,$$

where

$$w_i := \int_a^b L_i(x) \, w(x) \, \mathrm{d}x \ .$$

The formula $I_n(f)$ has degree of precision exactly $2n - 1$, which means that $I_n(x^k) = I(x^k)$ for $k = 0, 1, \ldots, 2n - 1$, but there is a polynomial Q of degree $2n$ for which $I_n(Q) \neq I(Q)$.

Proof: We will prove that the quadrature formula

$$I_n(f) = \sum_{i=1}^{n} w_i \, f(x_i)$$

(where the weights w_i are given by the formula above) has degree of precision exactly $2n - 1$ in three steps:

<u>Step 1</u>: *The degree of precision of the quadrature formula is $\geq n - 1$.*

Let V_{n-1} stand for the linear space of all polynomials of degree not exceeding $n-1$ defined for $x \in [a, b]$. We need to prove that our quadrature formula is exact for any $R \in V_{n-1}$.

The polynomials $L_i(x)$, $i = 1, 2, \ldots, n$ are Lagrange polynomials of degree $n-1$ such that $L_i(x_k) = \delta_{ik}$ for all $k = 1, 2, \ldots, n$. This implies that

$$\sum_{i=1}^{n} R(x_i) \, L_i(x)$$

is the Lagrange polynomial of degree $n-1$ that interpolates the polynomial R at the n points $x_1, x_2, \ldots, x_n$. Recalling that R is a polynomial of degree of degree $n-1$, we conclude that

$$R(x) = \sum_{i=1}^{n} R(x_i) \, L_i(x) \qquad \forall \, x \in [a, b] \ .$$

Then we have

$$
\begin{aligned}
I(R) \;=\; & \int_a^b R(x) \, w(x) \, \mathrm{d}x = \int_a^b \left[\sum_{i=1}^{n} R(x_i) \, L_i(x) \right] w(x) \, \mathrm{d}x \\
\;=\; & \sum_{i=1}^{n} R(x_i) \left[\int_a^b L_i(x) \, w(x) \, \mathrm{d}x \right] = \sum_{i=1}^{n} R(x_i) \, w_i = I_n(R) \ ,
\end{aligned}
$$

therefore the quadrature formula is exact for any polynomial R of degree up to and including $n-1$, or, in other words, the degree of precision of the quadrature formula is $\geq n-1$.

Step 2: *The degree of precision of the quadrature formula is $\geq 2n-1$.*

Let P be a polynomial of degree $\leq 2n-1$, i.e., $P \in V_{2n-1}$. Divide P by G_n (which is a polynomial of degree n to obtain

$$P(x) = Q(x) \, G_n(x) + R(x) \ ,$$

where $Q \in V_{n-1}$ and $R \in V_{n-1}$. Since the polynomials $G_0, G_1, G_2, \ldots, G_{n-1}$ form a basis of the linear space V_{n-1}, we can write Q in the form

$$Q(x) = \sum_{i=0}^{n-1} q_i \, G_i(x) \ .$$

Then

$$
\begin{aligned}
I(P) \;=\; & I(Q \, G_n + R) = \int_a^b \left[Q(x) \, G_n(x) + R(x) \right] w(x) \, \mathrm{d}x \\
\;=\; & \int_a^b \left[\sum_{i=0}^{n-1} q_i \, G_i(x) \, G_n(x) \right] w(x) \, \mathrm{d}x + \int_a^b R(x) \, w(x) \, \mathrm{d}x \\
\;=\; & 0 + \int_a^b R(x) \, w(x) \, \mathrm{d}x = I(R) \ ,
\end{aligned}
$$

6

because all polynomials G_0, G_1, ..., G_{n-1} are orthogonal to G_n (with respect to the inner product defined with the weight function w):

$$\langle G_i, G_n \rangle = \int_a^b G_i(x)\, G_n(x)\, w(x)\, \mathrm{d}x = 0 \qquad \forall\, i = 0, 1, 2, \ldots, n-1 \ .$$

Now recall that R is a polynomial of degree $\leq n-1$ to conclude that $I(R) = I_n(R)$ according to what we proved in Step 1.

Since x_i are roots of the polynomial G_n, it follows that

$$P(x_i) = Q(x_i)\, G_n(x_i) + R(x_i) = 0 + R(x_i) = R(x_i) \ ,$$

which implies that $I_n(R) = I_n(P)$. Combining these results, we obtain

$$I(P) = I_n(P) \qquad \forall\, P \in V_{2n-1} \ .$$

Step 3: *The degree of precision of the quadrature formula is $2n - 1$.*

We already know that the degree of precision is at least $2n - 1$, so to prove that it is *exactly* $2n - 1$, it is enough to find *one* polynomial of degree $2n$ for which the quadrature formula is *not* exact. Note that the polynomial $G_n^2(x)$ has degree exactly $2n$, and that

$$I(G_n^2) = \int_a^b G_n^2(x)\, w(x)\, \mathrm{d}x > 0 \tag{2}$$

because the integrand, $G_n^2(x)\, w(x)$, is nonnegative and can be zero only at isolated points (a polynomial like G_n^2 has at most $2n$ roots, and w can vanish only at isolated points by the definition of weight function).

But, since x_i are roots of G_n for $i = 1, 2, \ldots, n$, the quadrature formula gives

$$I_n(G_n^2) = \sum_{i=1}^n w_i\, G_n^2(x_i) = 0 \ . \tag{3}$$

Comparing (2) and (3), we see that the quadrature formula is *not* exact for G_n^2. ∎

In the proof of the above theorem, we implicitly used that G_n has exactly n *simple* roots, all of which are inside the interval (a, b). This property is established rigorously in the following lemma.

Lemma 1. *Let w be a weight function on $[a, b]$, let n be a positive integer, and let G_0, G_1, ..., G_n be a family of polynomials such that the degree of G_k is equal to k for each $k = 0, 1, \ldots, n$, and $\langle G_i, G_j \rangle = 0$ if $i \neq j$, where the inner product is defined by (1).*

Then for each $k = 0, 1, 2, \ldots, n$, the polynomial G_k has exactly k real roots which are simple and lie in the interval (a, b).

Proof: First we will prove the existence of real roots for G_k, and then we will count those roots.

Step 1: *Existence of real roots of G_k.*
Since the degree of G_0 is zero, it follows that $G_0(x) = \alpha \; \forall \; x \in [a, b]$ for some constant $\alpha \neq 0$. (If you don't understand why $\alpha \neq 0$, look at page 87 of the book.) Then, for $k \geq 1$,

$$0 = \int_a^b G_0(x) \, G_k(x) \, w(x) \, \mathrm{d}x = \alpha \int_a^b G_k(x) \, w(x) \, \mathrm{d}x \; ,$$

where the first equality holds due to orthogonality. However, since w is positive on $[a, b]$ except possibly being equal to 0 at isolated points, and the integral $\int_a^b G_k(x) \, w(x) \, \mathrm{d}x$ is equal to zero, we conclude that the function G_k must change sign in (a, b). Since G_k is a continuous function (all polynomials are continuous functions), the Intermediate Value Theorem guarantees that G_k has a real root somewhere in (a, b).

Step 2: *Counting the roots of G_k.*
Suppose that G_k changes sign at exactly j points r_1, r_2, ..., r_j in (a, b) such that

$$a < r_1 < r_2 < \cdots < r_j < b \; .$$

Without loss of generality, assume that $G_k(x) > 0$ for $x \in (a, r_1)$. Then G_k alternates sign on (r_1, r_2), (r_2, r_3), ..., (r_j, b). Define the auxiliary function

$$Z(x) \; := \; (-1)^j \prod_{i=1}^{j} (x - r_j) \; \in \; V_j \; .$$

By construction, $Z(x)$ and $G_k(x)$ have the same sign for all $x \in [a, b]$. From this it follows that

$$\int_a^b Z(x) \, G_k(x) \, w(x) \, \mathrm{d}x > 0 \; . \tag{4}$$

Suppose that $j < k$. Since the polynomials G_0, G_1, G_1, ..., G_j form a basis for V_j, there exist constants z_0, z_1, z_2, ..., z_j such that

$$Z(x) = \sum_{i=0}^{j} z_i \, G_j(x) \; .$$

Substituting this expression into (4), we obtain

$$\int_a^b Z(x) \, G_k(x) \, w(x) \, \mathrm{d}x \; = \; \int_a^b \left[\sum_{i=0}^{j} z_i \, G_j(x) \right] G_k(x) \, w(x) \, \mathrm{d}x$$

$$= \; \sum_{i=0}^{j} z_i \int_a^b G_j(x) \, G_k(x) \, w(x) \, \mathrm{d}x = 0 \; ,$$

8

where we have used that $j < k$, which implies that the polynomials G_j and G_k are orthogonal. The last equality clearly contradicts (4)! Hence the assumption that $j < k$ was wrong, i.e., we must have that $j \geq k$. However, since the degree of the polynomial G_k is k, G_k cannot have more than k roots, which implies that $j = k$. ∎

Example: As an example of application of the above Theorem, let us rederive the quadrature formula we derived in class, namely,

$$\int_{-1}^{1} f(x)\, \mathrm{d}x \approx f\left(-\frac{1}{\sqrt{3}}\right) + f\left(\frac{1}{\sqrt{3}}\right) \ .$$

This formula can be obtained by using the Legendre polynomials P_k defined on page 3. Take the polynomials $P_0(x) = 1$, $P_1(x) = x$, and $P_2(x) = x^2 - \frac{1}{3}$, which are an orthogonal basis in the linear space $V_2(-1, 1; w(x) \equiv 1)$. Since $n = 2$, we expect that the formula that we will obtain will have degree of precision $2n - 1 = 2(2) - 1 = 3$. As in the Theorem above, define $x_1 = -\frac{1}{\sqrt{3}}$ and $x_2 = \frac{1}{\sqrt{3}}$ to be the zeros of the polynomial P_2, define the polynomials (of degree $n - 1 = 2 - 1 = 1$)

$$L_1(x) = \frac{x - x_2}{x_1 - x_2} \ , \qquad L_2(x) = \frac{x - x_1}{x_2 - x_1} \ ,$$

and the weights

$$w_1 = \int_{-1}^{1} L_1(x)\, w(x)\, \mathrm{d}x = \int_{-1}^{1} \frac{x - x_2}{x_1 - x_2} 1\, \mathrm{d}x = \frac{1}{x_1 - x_2} \left(\frac{x^2}{2} - x_2 x\right)\Big|_{x=-1}^{1} = 1 \ ,$$

and, similarly, $w_2 = \int_{-1}^{1} L_2(x)\, w(x)\, \mathrm{d}x = \cdots = 1$, to obtain

$$I_2(f) = \sum_{i=1}^{2} w_i\, f(x_i) = f\left(-\frac{1}{\sqrt{3}}\right) + f\left(\frac{1}{\sqrt{3}}\right) \ .$$

As an exercise, check that the degree of accuracy of this quadrature formula is indeed 3.