Orthogonal polynomials and Gaussian quadrature.

The notation in these notes is not the standard notation, but is close to what we used in class. The subject of orthogonal polynomials is very large and what is here does not even start to give an idea about the general theory. The wikipedia page
`https://en.wikipedia.org/wiki/Orthogonal_polynomials`
is not a bad place to get more information.

In class with proved:

Proposition 1. *Let $(H, \langle\ ,\ \rangle)$ be an inner product space and V a finite dimensional subspace of H with an orthonormal basis $e_1, e_2, \ldots, e_n$. Let $f \in H$. Then the element v of V that minimizes $\|f - v\|$ is*

$$v_0 = \sum_{j=1}^{n} \langle f, e_j \rangle e_j. \qquad \square$$

A natural set up for this is $H = L^2([a, b])$ with the inner product

$$\langle f, g \rangle = \int_a^b f(x)g(x)\, dx$$

and $V = \mathcal{P}_n$ is the space of polynomials of degree $\le n$. So we would like to find an orthonormal basis of $\mathcal{P}_n$. Of course one way to do this would be to apply the Gram-Schmidt procedure to the basis $1, x, x^2, \ldots, x^{n-1}, x^n$ of $\mathcal{P}_n$, but that is a miss. Here is anther method.

Problem 1. This is likely something you have seen and so likely is review. Let f and g be real valued functions on $\mathbb{R}$ (or an interval in $\mathbb{R}$). Show

$$\frac{d^n}{dx^n}(f(x)g(x)) = \sum_{k=0}^{n} \binom{n}{k} f^{(k)}(x) g^{(n-k)}(x)$$

where $f^{(0)} = f$ and for $k \ge 1$ the function $f^{(k)}$ is the k-th derivative of f. $\square$

We now fix an interval $[a, b] \subseteq \mathbb{R}$.

Problem 2. Use the previous problem to show that if $m < n$ and

$$H(x) = \frac{d^m}{dx^m}\big((x-a)^n (x-b)^n\big)$$

then

$$H(a) = H(b) = 0. \qquad \square$$

Let $P_n(x)$ be the polynomial

$$P_n(x) = \frac{d^n}{dx^n}\big((x-a)^n (x-b)^n\big)$$

with the convention $P_0(x) = 1$.

Problem 3. Show that $P_n(x)$ is a polynomial of degree n and that if $f(x)$ is n times differentiable on $[a,b]$, then

$$\int_a^b f(x)P_n(x)\,dx = (-1)^n \int_a^b f^{(n)}(x)(x-a)^n(x-b)^n\,dx.$$

Hint: Integration by parts. □

Problem 4. Show that if $h(x)$ is a polynomial with $\deg(h(x)) < n$, then

$$\int_a^b h(x)P_n(x)\,dx = 0.$$

Thus the sequence of polynomials $P_0(x), P_1(x), P_2(x), \ldots$ are mutually orthogonal:

$$\int_a^b P_m(x)P_n(x)\,dx = 0 \qquad \text{for } m \neq n.$$ □

While this sequence of polynomials is orthogonal, it is not normalized. So we still have to compute

$$\|P_n\|_{L^2}^2 = \int_a^b P_n(x)^2\,dx.$$

We do this in steps.

Problem 5. Show

$$\begin{aligned}\int_a^b P_n(x)^2\,dx &= \int_a^b P_n(x)P_n(x)\,dx \\ &= (-1)^n \int_a^b P_n^{(n)}(x)(x-a)^n(x-b)^n\,dx \\ &= (-1)^n(2n)! \int_a^b (x-a)^n(x-b)^n\,dx.\end{aligned}$$

Hint: Problem 3. □

Thus if we set

$$I(m,n) = \int_a^b (x-a)^m(x-b)^n\,dx$$

we have reduce our problem to computing $I(n,n)$.

Problem 6. If $m \geq 1$ use integration do a change of variable to show

$$I(n,m) = (-1)^{m+n} I(m,n)$$

and integration by parts to show for $m \geq 1$ that

$$I(m,n) = -\frac{m}{n+1} I(m-1,n+1).$$ □

Problem 7. Now show that if $m \le n$

$$\begin{aligned} I(m,n) &= (-1)^m \frac{m!n!}{(m+n)!} I(0, m+n) \\ &= (-1)^m \frac{m!n!}{(m+n)!} \int_a^b (x-b)^{m+n}\, dx \\ &= (-1)^n \frac{m!n!}{(m+n+1)!} (b-a)^{m+n+1} \end{aligned}$$

and therefore

$$I(n,n) = (-1)^n \frac{(n!)^2}{(2n+1)!} (b-a)^{2n+1}. \qquad \square$$

Problem 8. Put the pieces together to get that

$$\int_a^b P_n(x)^2\, dx = \frac{(n!)^2}{2n+1} (b-a)^{2n+1}$$

and therefore if

$$c_n = \frac{n!}{\sqrt{2n+1}} (b-a)^{n+\frac{1}{2}}$$

and

$$Q_n(x) = \frac{P_n(x)}{c_n}$$

then $Q_0(x), Q_1(x), Q_2(x), \ldots$ is an orthonormal sequence of polynomials in $L^2([a,b])$ and $Q_0(x), Q_1(x), \ldots, Q_n(x)$ is an orthonormal basis of $\mathcal{P}_n$. $\square$

Let us now prove some properties of the polynomials $P_n(x)$.

Proposition 2. *The roots of $P_n(x)$ are all real, distanti, and in the open interval (a,b).*

Problem 9. Prove this. *Hint:* If this is false, then $P_n(x)$ will change sign at most $(n-1)$ times in (a,b). Let $r_1, r_2, \ldots, r_k \in (a,b)$ be the points where $P_n(x)$ changes sign in (a,b). Let

$$h(x) = (x-r_1)(x-r_2)\cdots(x-r_k).$$

As $k < n$ this polynomial is orthogonal to $P_n(x)$ and therefore

$$\int_a^b h(x) P_n(x)\, dx = 0.$$

But $P_n(x)$ and $h(x)$ change signs at the same points in (a,b) and therefore, except at the zeros of $P_n(x)$, the function $h(x)P_n(x)$ is either always positive or always negative on (a,b) which leads to a contradiction. $\square$

Our next goal is to choose points $r_1, r_2, \ldots, r_n \in (a,b)$ and constants $w_1, w_1, \ldots, w_n$ so that for a continuous function the sum

$$\sum_{k=1}^n w_k f(r_k) \approx \int_a^b f(x)\, dx.$$

That is we can find points $r_1, r_2, \ldots, r_n$ and weights $w_1, w_2, \ldots, w_n$ so that if we sample the function $f(x)$ at the points $r_1, r_2, \ldots, r_n$ and take a weighted average at these points, then the weighted average is a good approximation to the integral. Our method of choosing the points and weights will be to choose them so that the formula is exact on as many polynomial of low degree as possible.

Our (well really Gauss's) choice of the points where we will be doing the sampling is

$$r_1, r_2, \ldots, r_n = \text{roots of } P_n(x) \text{ listed in increasing order.}$$

To get the weights we need the ***Lagrange interpolation polynomials***. These are the polynomials $L_1(x), L_2(x), \ldots, L_n(x)$ such that

$$L_i(r_j) = \delta_{ij} = \begin{cases} 1, & i = j; \\ 0, & i \neq j. \end{cases}$$

Here is an explicit formula for these:

$$L_j(x) = \prod_{k \in \{1,2,\ldots,n\} \setminus \{j\}} \frac{x - r_k}{r_j - r_k}.$$

So for $n = 4$ these are

$$\begin{aligned} L_1(x) &= \frac{(x - r_2)(x - r_3)(x - r_4)}{(r_1 - r_2)(r_1 - r_3)(r_1 - r_4)}, \\ L_2(x) &= \frac{(x - r_1)(x - r_3)(x - r_4)}{(r_2 - r_1)(r_2 - r_3)(r_2 - r_4)}, \\ L_3(x) &= \frac{(x - r_1)(x - r_2)(x - r_4)}{(r_3 - r_1)(r_3 - r_2)(r_3 - r_4)}, \\ L_4(x) &= \frac{(x - r_1)(x - r_2)(x - r_3)}{(r_4 - r_1)(r_4 - r_2)(r_4 - r_3)}. \end{aligned}$$

Problem 10. Let $c_1, c_2, \ldots, c_n$ be any real numbers. Show that

$$h(x) = \sum_{k=1}^{n} c_k L_k(x)$$

is the unique polynomial of degree at most $(n - 1)$ with

$$h(r_j) = c_j$$

for $j = 1, 2, \ldots, n$. □

We now define the weights

$$w_k := \int_a^b L_j(x)\, dx.$$

Problem 11. Show

$$\sum_{k=1}^{n} w_k = (b - a).$$

Hint: Use $c_k = 1$ in Problem 10 to conclude

$$\sum_{k=1}^{n} L_k(x) \equiv 1$$

and integrate this over the interval $[a, b]$. $\square$

Theorem 3 (Gauss Quadrature Formula). *For any polynomial $f(x)$ with degree $\leq 2n-1$*

$$\sum_{k=1}^{n} w_k f(r_k) = \int_a^b f(x)\,dx.$$

Problem 12. Prove this. *Hint:* Use the division algorithm to divide the polynomial $P_n(x)$ into $f(x)$ so that

$$f(x) = q(x)P_n(x) + r(x) \qquad \text{where } \deg(g(x)) < n, \text{ or } g(x) \equiv 0.$$

Since $\deg(P_n(x)) = n$ and $\deg(f(x)) \leq 2n-1$ we see $\deg(q(x)) < n$. Therefore by Problem 4

$$\int_a^b q(x)P_n(x)\,dx = 0.$$

Use this to show

$$\int_a^b f(x)\,dx = \int_a^b g(x)\,dx.$$

As the numbers $r_1, r_2, \ldots, r_n$ are the zeros of we have

$$f(r_k) = g(r_k).$$

Therefore by Problem 10

$$g(x) = \sum_{k=1}^{n} f(r_k)L_k(x)$$

and integration gives

$$\int_a^b g(x)\,dx = \sum_{k=1}^{n} f(r_k)\int_a^b L_k(x)\,dx.$$

Combining these formulas and using the definition of the w_k's gives the result. $\square$

Corollary 4. *Gauss was very clever.* $\square$