



# Gaussian Quadrature & Tridiagonal Matrices

**Kyle Monette**

Department of Mathematics and Applied Mathematical Sciences  
University of Rhode Island

Applied Mathematics and Scientific Computing Seminar

December 9, 2024

# Introduction

- 1 Introduce Quadrature Rules.
- 2 Generalize this to *Gaussian* Quadrature rules.
- 3 Illustrate some examples.
- 4 Introduce the tridiagonal matrix eigenvalue problem.

Next time:

- 1 How can the weights from Gaussian Quadrature be used for other linear algebra algorithms?
- 2 Lanczos algorithm, Ritz Values & Vectors , ...

# Problem Statement

**Goal:** Given a function  $f \in C[a, b]$ , distinct  $\{x_0, x_1, \dots, x_n\} \subset [a, b]$ , and  $\{w_0, w_1, \dots, w_n\} \in \mathbb{R}$ , we wish to write

$$\int_a^b f(x) \, dx \approx \sum_{j=0}^n w_j f(x_j).$$

**Motivation:**

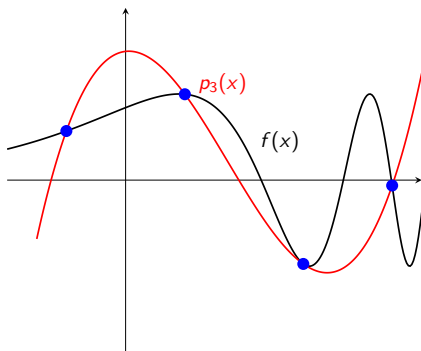
- Perhaps we only know  $f$  by its action on some points.
- There may not be an elementary antiderivative of  $f$ .

# Introduction

Our discussion will be based on **interpolary** quadrature methods.

Find the  $n^{\text{th}}$  degree polynomial **interpolant**  $p_n$  to  $f$  at the nodes  $x_0, \dots, x_n$ .

$$\int_a^b f(x) dx \approx \int_a^b p_n(x) dx.$$



If  $f \in C^{n+1}[a, b]$ , the error is

$$\frac{1}{(n+1)!} \int_a^b \prod_{i=0}^n (x - x_i) f^{(n+1)}(\xi(x)) dx$$

# Constructing $p_n(x)$

Here and going forward, we let

$$\mathcal{P}_n = \{\text{polynomials of degree at most } n\}.$$

Given  $n + 1$  distinct points  $\{x_0, \dots, x_n\} \subset [a, b]$ , the **Lagrange basis** for  $\mathcal{P}_n$  is

$$\{\ell_0, \ell_1, \dots, \ell_n\}, \quad \text{where} \quad \ell_j(x) = \prod_{\substack{k=0 \\ k \neq j}}^n \frac{x - x_k}{x_j - x_k}.$$

Then we have the **unique** degree  $n$  polynomial interpolant to  $f$ :

$$p_n(x) = \sum_{j=0}^n f(x_j) \ell_j(x)$$

# Determining Weights

Now, notice that

$$\begin{aligned}\int_a^b f(x) \, dx &\approx \int_a^b p_n(x) \, dx \\ &= \int_a^b \sum_{j=0}^n f(x_j) \ell_j(x) \, dx \\ &= \sum_{j=0}^n f(x_j) \int_a^b \ell_j(x) \, dx\end{aligned}$$

Let  $w_j = \int_a^b \ell_j(x) \, dx$  for  $j = 0, \dots, n$ . Then

$$\int_a^b f(x) \, dx \approx \sum_{j=0}^n w_j f(x_j).$$

# Quadrature Rule

$$\int_a^b f(x) dx \approx \sum_{j=0}^n w_j f(x_j), \quad w_j = \int_a^b \ell_j(x) dx \quad (1)$$

**Note:** When  $f \in \mathcal{P}_n$ , we have equality.

## Definition 1 (Degree of Exactness)

The quadrature rule (1) is said to have *degree of exactness*  $m$  if it exactly integrates all polynomials in  $\mathcal{P}_m$ .

So the  $n + 1$  point quadrature rule (1) has degree of exactness  $n$ .

We will raise the degree of exactness in a clever way.

# Generalization of the Problem

To increase the degree of exactness, we consider a **generalized problem**:

$$\int_a^b f(x)w(x) dx \approx \sum_{j=0}^n w_j f(x_j)$$

for a weight **function**  $w(x) \in C[a, b]$  and  $w(x) \geq 0$  on  $[a, b]$ .

$$\begin{aligned} \int_a^b f(x)w(x) dx &\approx \int_a^b p_n(x)w(x) dx \\ &= \int_a^b \sum_{j=0}^n f(x_j)\ell_j(x)w(x) dx \\ &= \sum_{j=0}^n f(x_j) \int_a^b \ell_j(x)w(x) dx \\ &= \sum_{j=0}^n w_j f(x_j) \end{aligned}$$

where  $w_j = \int_a^b \ell_j(x)w(x) dx$ .



# Generalization of the Problem

This yields the following generalized  $n + 1$  point quadrature rule:

$$\int_a^b f(x)w(x) dx \approx \sum_{j=0}^n f(x_j)w_j \quad \text{where} \quad w_j = \int_a^b \ell_j(x)w(x) dx \quad (2)$$

(Still has degree of exactness  $n$ )

Choosing the nodes  $x_j$  and weights  $w_j$  in a clever way, we can dramatically increase the degree of exactness of (2) for any  $w(x)$ .

To do so, we need to use orthogonal polynomials.

# Orthogonal Polynomials

As before, consider a weight function  $w(x) \in C[a, b]$  where  $w(x) \geq 0$  and  $w(x) \neq 0$  on any subinterval of  $[a, b]$ .

## Definition 2 (Orthogonal Polynomials)

A set of polynomials  $\{\phi_0, \dots, \phi_{n+1}\}$  (where  $\deg(\phi_k) = k$ ) is said to be an **orthogonal set of polynomials** with respect to the **weight function**  $w(x)$  if

$$\langle \phi_i(x), \phi_j(x) \rangle := \int_a^b \phi_i(x) \phi_j(x) w(x) dx = 0 \quad \forall i \neq j.$$

Orthonormal polynomials (i.e.,  $\langle \phi_i, \phi_i \rangle = 1$ ) are not necessary here.

We will now show that the  $n + 1$  point generalized quadrature rule has degree of exactness  $2n + 1$ .

# Generalized Rule Degree of Exactness $2n + 1$

Let  $p$  be degree  $2n + 1$  and  $\{\phi_0, \dots, \phi_{n+1}\}$  a set of orthogonal polynomials. Then there exists  $q, r \in \mathcal{P}_n$  such that

$$p(x) = \phi_{n+1}(x)q(x) + r(x).$$

Then

$$\begin{aligned} \int_a^b p(x)w(x) dx &= \int_a^b \phi_{n+1}(x)q(x)w(x) dx + \int_a^b r(x)w(x) dx \\ &= \langle \phi_{n+1}(x), q(x) \rangle + \int_a^b r(x)w(x) dx \\ &= \left\langle \phi_{n+1}(x), \sum_{i=0}^n c_i \phi_i(x) \right\rangle + \int_a^b r(x)w(x) dx \\ &= \int_a^b r(x)w(x) dx \end{aligned}$$

# Generalized Rule Degree of Exactness $2n + 1$

$$p(x) = \phi_{n+1}(x)q(x) + r(x), \quad \int_a^b p(x)w(x) dx = \int_a^b r(x)w(x) dx.$$

Apply the  $n + 1$  point quadrature rule to  $p$ . (recall:  $\deg(p) = 2n + 1$ )

$$\begin{aligned} \int_a^b p(x)w(x) dx &\approx \sum_{j=0}^n w_j p(x_j) \\ &= \sum_{j=0}^n w_j \phi_{n+1}(x_j) q(x_j) + \sum_{j=0}^n w_j r(x_j) \\ &= \sum_{j=0}^n w_j \phi_{n+1}(x_j) q(x_j) + \int_a^b r(x)w(x) dx \end{aligned}$$

If this sum is always 0, then

$$\sum_{j=0}^n w_j p(x_j) = \int_a^b r(x)w(x) dx = \int_a^b p(x)w(x) dx.$$

# Generalized Rule Degree of Exactness $2n + 1$

If we make

$$\sum_{j=0}^n w_j \phi_{n+1}(x_j) q(x_j) = 0$$

for any choice of nodes and weights, then the generalized  $n + 1$  point quadrature rule

$$\int_a^b f(x) w(x) dx \approx \sum_{j=0}^n f(x_j) w_j \quad \text{where} \quad w_j = \int_a^b \ell_j(x) w(x) dx$$

has degree of exactness  $2n + 1$  because it is exact for all  $p \in \mathcal{P}_{2n+1}$ .

So, we choose the nodes  $\{x_0, \dots, x_n\}$  to be the **roots of  $\phi_{n+1}$** .

# Gaussian Quadrature

## Theorem 3

Let  $\{\phi_0, \dots, \phi_{n+1}\}$  be a set of orthogonal polynomials on  $[a, b]$  with respect to the weight function  $w(x)$ . Then  $\phi_k$  has  $k$  distinct zeros in  $[a, b]$  for each  $k = 0, 1, \dots, n+1$ .

## Theorem 4 (Gaussian Quadrature)

Given  $n$ , an interval  $[a, b]$ , and a weight function  $w(x)$ , choose the nodes  $x_i$  to be the  $n+1$  roots of  $\phi_{n+1}$  and the weights  $w_i$  to be  $w_i = \int_a^b \ell_i(x) w(x) dx$ . Then the quadrature rule

$$\int_a^b f(x) w(x) dx \approx \sum_{j=0}^n w_j f(x_j)$$

is *exact* for all  $f \in \mathcal{P}_{2n+1}$ .

**Observation:** The nodes and weights do not depend on  $f$ !!

# Generating Orthogonal Polynomials

How are the orthogonal polynomials  $\{\phi_0, \dots, \phi_{n+1}\}$  constructed?

**Answer:** Gram-Schmidt Algorithm, starting with  $\phi_0 := 1$ .

One can show the **three-term recurrence**:

$$\phi_0 = 1$$

$$\phi_1 = x - \frac{\langle x, 1 \rangle}{\langle 1, 1 \rangle}$$

$$\phi_k = x\phi_{k-1} - \frac{\langle x\phi_{k-1}, \phi_{k-1} \rangle}{\langle \phi_{k-1}, \phi_{k-1} \rangle} \phi_{k-1} - \frac{\langle x\phi_{k-1}, \phi_{k-2} \rangle}{\langle \phi_{k-2}, \phi_{k-2} \rangle} \phi_{k-2}, \quad k \geq 2$$

# Generating Orthogonal Polynomials

By shifting indices, then

$$\begin{aligned}\phi_{-1} &= 0 \\ \phi_0 &= 1 \\ \phi_{k+1} &= x\phi_k - \alpha_k\phi_k - \beta_k\phi_{k-1}, \quad k \geq 0\end{aligned}$$

where we define

$$\alpha_k = \frac{\langle x\phi_k, \phi_k \rangle}{\langle \phi_k, \phi_k \rangle} \quad \text{for } k \geq 0, \quad \beta_k = \frac{\langle x\phi_k, \phi_{k-1} \rangle}{\langle \phi_{k-1}, \phi_{k-1} \rangle} \quad \text{for } k \geq 1.$$

Note that this recurrence generates **monic** polynomials.



# Examples of Orthogonal Polynomials

For each weight function  $w(x)$  there exists a system of orthogonal polynomials.

Weight	Interval	Orthogonal Polynomials
$w(x) = 1$	$[-1, 1]$	Legendre
$w(x) = \frac{1}{\sqrt{1-x^2}}$	$(-1, 1)$	Chebyshev
$w(x) = e^{-x}$	$(0, \infty)$	Laguerre
$w(x) = e^{-x^2}$	$(-\infty, \infty)$	Hermite

# Example: Legendre Polynomials

When  $w(x) = 1$  and  $[a, b] = [-1, 1]$ , we get the **Legendre Polynomials**.

One can show that

$$\alpha_k = \frac{\langle x\phi_k, \phi_k \rangle}{\langle \phi_k, \phi_k \rangle} = 0 \quad 0 \leq k \leq n, \quad \beta_k = \frac{\langle x\phi_k, \phi_{k-1} \rangle}{\langle \phi_{k-1}, \phi_{k-1} \rangle} = \frac{k^2}{4k^2 - 1} \quad 1 \leq k \leq n$$

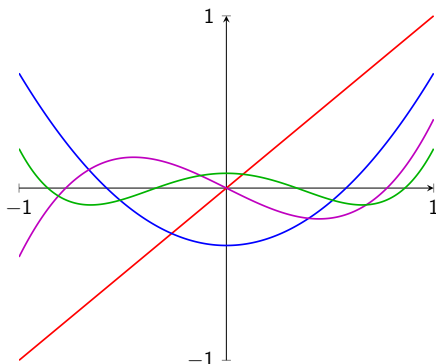
$$\phi_0(x) = 1$$

$$\phi_1(x) = x$$

$$\phi_2(x) = x^2 - \frac{1}{3}$$

$$\phi_3(x) = x^3 - \frac{3}{5}x$$

$$\phi_4(x) = x^4 - \frac{6}{7}x^2 + \frac{3}{35}$$



# Example: Chebyshev Polynomials

When  $w(x) = \frac{1}{\sqrt{1-x^2}}$  and  $[a, b] = (-1, 1)$ , we get the **Chebyshev Polynomials**.

$$\alpha_k = 0 \quad 0 \leq k \leq n, \quad \beta_k = \begin{cases} 1/2 & k = 1 \\ 1/4 & 2 \leq k \leq n \end{cases}, \quad w_j = \frac{\pi}{n+1} \quad 0 \leq j \leq n$$

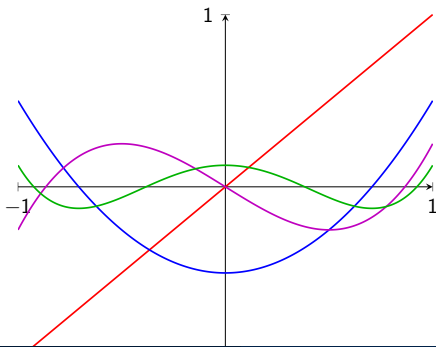
$$\phi_0(x) = 1$$

$$\phi_1(x) = x$$

$$\phi_2(x) = x^2 - \frac{1}{2}$$

$$\phi_3(x) = x^3 - \frac{3}{4}x$$

$$\phi_4(x) = x^4 - x^2 + \frac{1}{8}$$



# Progress So Far

The orthogonal polynomials  $\{\phi_0, \dots, \phi_{n+1}\}$  can be generated cheaply via a three-term recurrence.

The nodes  $\{x_0, \dots, x_n\}$  can be chosen to be the roots of  $\phi_{n+1}$ , and the weights  $\{w_0, \dots, w_n\}$  are found by  $w_j = \int_a^b \ell_j(x) w(x) dx$ .

Is there a way to more easily compute the nodes and weights?

# The Tridiagonal Connection

From

$$\alpha_k = \frac{\langle x\phi_k, \phi_k \rangle}{\langle \phi_k, \phi_k \rangle} \quad \text{for } k \geq 0, \quad \beta_k = \frac{\langle x\phi_k, \phi_{k-1} \rangle}{\langle \phi_{k-1}, \phi_{k-1} \rangle} \quad \text{for } k \geq 1$$

we define

$$\phi(x) = \begin{bmatrix} \phi_0(x) \\ \phi_1(x) \\ \vdots \\ \phi_n(x) \end{bmatrix}_{(n+1) \times 1} \quad T = \begin{bmatrix} \alpha_0 & 1 & & & \\ \beta_1 & \alpha_1 & 1 & & \\ & \beta_2 & \ddots & \ddots & \\ & & \ddots & \alpha_{n-1} & 1 \\ & & & \beta_n & \alpha_n \end{bmatrix}_{(n+1) \times (n+1)}.$$

Then

$$x\phi(x) = T\phi(x) + \phi_{n+1}(x)e_{n+1}.$$

# The Tridiagonal Connection

$$x\phi(x) = T\phi(x) + \phi_{n+1}(x)e_{n+1}$$

Suppose  $\phi_{n+1}(\lambda) = 0$ . Then

$$\lambda\phi(\lambda) = T\phi(\lambda)$$

so  $(\lambda, \phi(\lambda))$  is an **eigenpair** of  $T$ .

Quadrature Nodes    =    Roots of  $\phi_{n+1}(x)$     =    Evals. of  $T$

# The Tridiagonal Connection

Using a similarity transformation with

$$D = \text{diag}\left(1, \sqrt{\beta_1}, \sqrt{\beta_1\beta_2}, \dots, \sqrt{\beta_1 \dots \beta_n}\right)$$

then

$$D^{-1}TD := J = \begin{bmatrix} \alpha_0 & \sqrt{\beta_1} & & & \\ \sqrt{\beta_1} & \alpha_1 & \sqrt{\beta_2} & & \\ & \sqrt{\beta_2} & \ddots & \ddots & \\ & & \ddots & \alpha_{n-1} & \sqrt{\beta_n} \\ & & & \sqrt{\beta_n} & \alpha_n \end{bmatrix}_{(n+1) \times (n+1)}.$$

If  $(\lambda, \phi(\lambda))$  is an eigenpair of  $T$ , then  $(\lambda, D^{-1}\phi(\lambda))$  is an eigenpair of  $J$ .

# Calculating Weights

Eigenpairs of  $J$  are of the form  $(\lambda, D^{-1}\phi(\lambda))$ :

$$v := D^{-1}\phi(\lambda) = \begin{bmatrix} \phi_0(\lambda) \\ \phi_1(\lambda)/\sqrt{\beta_1} \\ \vdots \\ \phi_n(\lambda)/\sqrt{\beta_1 \dots \beta_n} \end{bmatrix}.$$

In 1969, Golub and Welsch proved that the weights  $w_i$  satisfy

$$w_i = \beta_0 \cdot \frac{v_{1i}^2}{\|v_i\|^2} \quad 0 \leq i \leq n$$

where  $\beta_0 = \langle 1, 1 \rangle = \int_a^b w(x) dx$ .



# Gaussian Quadrature

Given  $f \in C[a, b]$ , weight function  $w(x)$ , and integer  $n$ , form the matrix  $J$

$$J = \begin{bmatrix} \alpha_0 & \sqrt{\beta_1} & & & \\ \sqrt{\beta_1} & \alpha_1 & \sqrt{\beta_2} & & \\ & \sqrt{\beta_2} & \ddots & \ddots & \\ & & \ddots & \alpha_{n-1} & \sqrt{\beta_n} \\ & & & \sqrt{\beta_n} & \alpha_n \end{bmatrix}$$

with values defined by the three-term recurrence. Find eigenpairs  $(\lambda_i, v_i)$  of  $J$ . By choosing the nodes  $x_i = \lambda_i$  and weights  $w_i = \beta_0 \cdot \frac{v_{1i}^2}{\|v_i\|^2}$  then

$$\int_a^b f(x) w(x) dx = \sum_{j=0}^n w_j f(x_j)$$

for all  $f \in \mathcal{P}_{2n+1}$ .

## Example: Gauss-Legendre Quadrature

Consider the interval  $[-1, 1]$  and  $w(x) = 1$ . Recall that this generates the (monic) Legendre polynomials

$$\phi_0(x) = 1, \quad \phi_1(x) = x, \quad \phi_2(x) = x^2 - \frac{1}{3}, \quad \dots$$

by using

$$\alpha_k = 0 \quad 0 \leq k \leq n, \quad \beta_k = \frac{k^2}{4k^2 - 1} \quad 1 \leq k \leq n$$

# Example: Gauss-Legendre Quadrature

For example, let  $n = 4$ . Then  $J$  is

$$J = \begin{bmatrix} 0 & 0.5774 & & & \\ 0.5774 & 0 & 0.5164 & & \\ & 0.5164 & 0 & 0.5071 & \\ & & 0.5071 & 0 & 0.5040 \\ & & & 0.5040 & 0 \end{bmatrix}.$$

Eigenvalues of  $J$  (nodes) and weights are

Nodes	-0.9062	-0.5385	0	0.5385	0.9062
Weights	0.2369	0.4786	0.5689	0.4786	0.2369

# Example: Gauss-Legendre Quadrature

With  $n = 4$ , we can **exactly** integrate polynomials of degree  $\leq 2n + 1 = 9$ :

$$\int_{-1}^1 x^9 + x^6 dx = \sum_{k=0}^4 w_j f(x_j) = 0.2857$$

$$\int_{-1}^1 x^{12} dx \quad \text{Error: } 0.008$$

$$\int_{-1}^1 \sin(e^{x^2}) dx \quad \text{Error: } 9.9 \times 10^{-4}$$

# Shifting to Other Intervals

Consider a rule on  $[c, d]$ , with nodes  $x_j$ , weights  $w_j$ , weight function  $w(x)$ .

We want to create a quadrature rule on  $[a, b]$ .

Define the affine transformation

$$\begin{aligned}\tau : [c, d] &\rightarrow [a, b], & \tau(x) &= a + \frac{b-a}{d-c}(x-c) \\ \tau^{-1} : [a, b] &\rightarrow [c, d], & \tau^{-1}(y) &= c + \frac{d-c}{b-a}(y-a)\end{aligned}$$

Note that

$$\begin{aligned}\int_a^b f(x)w(x) dx &= \int_{\tau^{-1}(a)}^{\tau^{-1}(b)} f(\tau(x))w(\tau(x))\tau'(x) dx \\ &= \frac{b-a}{d-c} \int_c^d f(\tau(x))w(\tau(x)) dx\end{aligned}$$

# Shifting to Other Intervals

$$\begin{aligned}\int_a^b f(x)w(x) dx &= \frac{b-a}{d-c} \int_c^d f(\tau(x))w(\tau(x)) dx \\ &= \frac{b-a}{d-c} \sum_{j=0}^n f(\tau(x_j)) w_j\end{aligned}$$

Then the nodes and weights for the  $[a, b]$  quadrature rule are

$$\hat{x}_j = \tau(x_j), \quad \hat{w}_j = \frac{b-a}{d-c} w_j$$

and the weight function is  $\hat{w}(x) = w(\tau(x))$ . That is,

$$\int_a^b f(x)\hat{w}(x) dx = \sum_{j=0}^n f(\hat{x}_j)\hat{w}_j.$$

# Tridiagonal Eigenvalue Problem

Computing nodes and weights requires forming and computing eigenpairs of

$$J = \begin{bmatrix} \alpha_0 & \sqrt{\beta_1} & & & \\ \sqrt{\beta_1} & \alpha_1 & \sqrt{\beta_2} & & \\ & \sqrt{\beta_2} & \ddots & \ddots & \\ & & \ddots & \alpha_{n-1} & \sqrt{\beta_n} \\ & & & \sqrt{\beta_n} & \alpha_n \end{bmatrix}$$

On the other hand, if we have an arbitrary symmetric tridiagonal matrix, what do quadrature weights of the matrix tell us?

# Next Time

- In numerical linear algebra, one may wish to approximate extreme eigenpairs of a large symmetric matrix  $A$  using an iterative method.
- In short: Given  $A \in \mathbb{R}^{n \times n}$  and  $m \ll n$ , then  $m$  extreme eigenpairs of  $A$  may be approximated by the eigenpairs of an  $m \times m$  symmetric tridiagonal matrix  $T$  (called Ritz pairs).
- We are interested in how the Gaussian Weights of  $T$  may accelerate convergence of the Ritz pairs.



Questions?

## Theorem 5

Let  $\{\phi_0, \dots, \phi_{n+1}\}$  be a set of orthogonal polynomials on  $[a, b]$  with respect to the weight function  $w(x)$ . Then  $\phi_k$  has  **$k$  distinct zeros** in  $[a, b]$  for each  $k = 0, 1, \dots, n+1$ .

Clearly the result holds for  $\phi_0$ . Suppose by way of contradiction that  $\phi_k$  changes sign on  $[a, b]$  at  $j$  distinct roots  $x_1, \dots, x_j$  where  $j < k$ . Define

$$q(x) = (x - x_1)(x - x_2) \dots (x - x_j) \in \mathcal{P}_j.$$

Note that  $q$  changes sign at  $x_1, \dots, x_j$  as well, so the product  $q(x)\phi_k(x)$  does not change sign on  $[a, b]$ . Because  $w(x) \geq 0$ , then  $q(x)\phi_k(x)w(x)$  does not change sign either on  $[a, b]$ . Because  $q \in \mathcal{P}_j$  and  $j < k$ , then

$$\langle \phi_k, q \rangle = \int_a^b \phi_k(x)q(x)w(x) dx = 0$$

This suggests that a continuous nonzero function which never changes sign on  $[a, b]$  has an integral that is zero, hence we have our contradiction. Therefore,  $\phi_k$  has *at least*  $k$  distinct zeros in  $[a, b]$ , so it has precisely  $k$ .

# Three-Term Recurrence

Suppose we have an orthogonal set  $\{\phi_0, \dots, \phi_{k-1}\}$  and want  $\phi_k$ . Note that  $x\phi_{k-1}$  has degree  $k$  and is not in the span of  $\{\phi_0, \dots, \phi_{k-1}\}$ . Orthogonalize  $x\phi_{k-1}$  against the previous  $\{\phi_0, \dots, \phi_{k-1}\}$  to get

$$\phi_k(x) = x\phi_{k-1}(x) - \sum_{j=0}^{k-1} \frac{\langle x\phi_{k-1}(x), \phi_j(x) \rangle}{\langle \phi_j(x), \phi_j(x) \rangle} \phi_j(x). \quad (3)$$

Now, notice that we can move the  $x$  to the other slot in the inner product above because

$$\begin{aligned} \langle x\phi_{k-1}(x), \phi_j(x) \rangle &= \int_a^b x\phi_{k-1}(x)\phi_j(x)w(x) dx \\ &= \int_a^b \phi_{k-1}(x)x\phi_j(x)w(x) dx \\ &= \langle \phi_{k-1}(x), x\phi_j(x) \rangle \end{aligned}$$

In this inner product, if  $j + 1$  (the degree of the polynomial in the second slot) is less than  $k - 1$  (the degree of the polynomial in the first slot), then by orthogonality the inner product is zero. That is,

$$j + 1 < k - 1 \quad \Rightarrow \quad \langle \phi_{k-1}(x), x\phi_j(x) \rangle = 0.$$

Equivalently if  $j < k - 2$  then the inner product is zero. This means that we can move the starting index of the summation in (3) up to  $k - 2$ :

$$\phi_k(x) = x\phi_{k-1}(x) - \sum_{j=k-2}^{k-1} \frac{\langle x\phi_{k-1}(x), \phi_j(x) \rangle}{\langle \phi_j(x), \phi_j(x) \rangle} \phi_j(x).$$