Numerical Analysis

Raphael Hauser
with thanks to Endre Süli

Oxford Mathematical Institute

HT 2019

Least-Squares Approximation

Consider the problem of least-squares approximation in the inner product space $\mathrm{L}^2_w(a,b)$, with inner product

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

and norm

$$\|f\|_2 = \langle f, f\rangle^{1/2}$$

where w is a weight-function, defined, positive and integrable on (a, b).

Theorem

If $f \in \mathrm{L}^2_w(a,b)$ *and* $p_n \in \Pi_n$ *is such that*

$$\langle f - p_n, r\rangle = 0 \qquad \forall r \in \Pi_n, \tag{12.1}$$

then

$$\|f - p_n\|_2 \leq \|f - r\|_2 \qquad \forall r \in \Pi_n,$$

i.e., p_n *is a best (weighted) least-squares approximation to* f *on* $[a,b]$.

Proof.

$$\|f - p_n\|_2^2$$

Proof.

$$\|f - p_n\|_2^2 = \langle f - p_n, f - p_n \rangle$$

Proof.

$$\begin{aligned}\|f - p_n\|_2^2 &= \langle f - p_n, f - p_n\rangle \\ &= \langle f - p_n, f - r\rangle + \langle f - p_n, r - p_n\rangle \qquad \forall r \in \Pi_n\end{aligned}$$

Proof.

$$
\begin{aligned}
\|f-p_n\|_2^2 &= \langle f-p_n, f-p_n\rangle \\
&= \langle f-p_n, f-r\rangle + \langle f-p_n, r-p_n\rangle \qquad \forall r\in\Pi_n
\end{aligned}
$$

Since $r-p_n\in\Pi_n$ the assumption (12.1) implies

Proof.

$$
\begin{aligned}
\|f - p_n\|_2^2 &= \langle f - p_n, f - p_n \rangle \\
&= \langle f - p_n, f - r \rangle + \langle f - p_n, r - p_n \rangle \qquad \forall r \in \Pi_n \\
&\ \text{Since } r - p_n \in \Pi_n \text{ the assumption (12.1) implies} \\
&= \langle f - p_n, f - r \rangle
\end{aligned}
$$

Proof.

$$
\begin{aligned}
\|f - p_n\|_2^2 &= \langle f - p_n, f - p_n \rangle \\
&= \langle f - p_n, f - r \rangle + \langle f - p_n, r - p_n \rangle \qquad \forall r \in \Pi_n \\
& \text{Since } r - p_n \in \Pi_n \text{ the assumption (12.1) implies} \\
&= \langle f - p_n, f - r \rangle \\
&\leq \|f - p_n\|_2 \|f - r\|_2 \text{ by the Cauchy–Schwarz inequality.}
\end{aligned}
$$

Proof.

$$
\begin{aligned}
\|f - p_n\|_2^2 &= \langle f - p_n, f - p_n \rangle \\
&= \langle f - p_n, f - r \rangle + \langle f - p_n, r - p_n \rangle \qquad \forall r \in \Pi_n \\
& \text{ Since } r - p_n \in \Pi_n \text{ the assumption (12.1) implies} \\
&= \langle f - p_n, f - r \rangle \\
&\leq \|f - p_n\|_2 \|f - r\|_2 \text{ by the Cauchy–Schwarz inequality.}
\end{aligned}
$$

Dividing both sides by $\|f - p_n\|_2$ gives the required result. □

Proof.

$$
\begin{aligned}
\|f - p_n\|_2^2 &= \langle f - p_n, f - p_n \rangle \\
&= \langle f - p_n, f - r \rangle + \langle f - p_n, r - p_n \rangle \qquad \forall r \in \Pi_n \\
&\quad \text{Since } r - p_n \in \Pi_n \text{ the assumption (12.1) implies} \\
&= \langle f - p_n, f - r \rangle \\
&\leq \|f - p_n\|_2 \|f - r\|_2 \text{ by the Cauchy–Schwarz inequality.}
\end{aligned}
$$

Dividing both sides by $\|f - p_n\|_2$ gives the required result. □

Remark: the converse is true too (see Problem Sheet 6, Q9).

This gives a direct way to calculate a best approximation:

This gives a direct way to calculate a best approximation: we want to find $p_n(x) = \sum_{k=0}^{n} \alpha_k x^k$ such that

$$\int_a^b w(x) \left(f - \sum_{k=0}^{n} \alpha_k x^k \right) x^i \, \mathrm{d}x = 0 \quad \text{for} \quad i = 0, 1, \ldots, n. \tag{12.2}$$

This gives a direct way to calculate a best approximation: we want to find $p_n(x) = \sum_{k=0}^{n} \alpha_k x^k$ such that

$$\int_a^b w(x) \left(f - \sum_{k=0}^{n} \alpha_k x^k \right) x^i \, \mathrm{d}x = 0 \quad \text{for} \quad i = 0, 1, \ldots, n. \tag{12.2}$$

[Note that (12.2) holds if, and only if,

$$\int_a^b w(x) \left(f - \sum_{k=0}^{n} \alpha_k x^k \right) \left(\sum_{i=0}^{n} \beta_i x^i \right) \mathrm{d}x = 0 \qquad \forall q = \sum_{i=0}^{n} \beta_i x^i \in \Pi_n.]$$

This gives a direct way to calculate a best approximation: we want to find $p_n(x) = \sum_{k=0}^{n} \alpha_k x^k$ such that

$$\int_a^b w(x) \left(f - \sum_{k=0}^{n} \alpha_k x^k \right) x^i \, \mathrm{d}x = 0 \quad \text{for} \quad i = 0, 1, \ldots, n. \tag{12.2}$$

[Note that (12.2) holds if, and only if,

$$\int_a^b w(x) \left(f - \sum_{k=0}^{n} \alpha_k x^k \right) \left(\sum_{i=0}^{n} \beta_i x^i \right) \mathrm{d}x = 0 \qquad \forall q = \sum_{i=0}^{n} \beta_i x^i \in \Pi_n.]$$

However, (12.2) implies that

$$\sum_{k=0}^{n} \left(\int_a^b w(x) x^{k+i} \, \mathrm{d}x \right) \alpha_k = \int_a^b w(x) f(x) x^i \, \mathrm{d}x \quad \text{for} \quad i = 0, 1, \ldots, n.$$

This is the component-wise statement of a matrix equation

$$A\alpha = \varphi, \tag{12.3}$$

to determine the coefficients $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_n)^{\mathrm{T}}$, where $A = \{a_{i,k}, i, k = 0, 1, \ldots, n\}$, $\varphi = (f_0, f_1, \ldots, f_n)^{\mathrm{T}}$,

$$a_{i,k} = \int_a^b w(x) x^{k+i} \, \mathrm{d}x \quad \text{and} \quad f_i = \int_a^b w(x) f(x) x^i \, \mathrm{d}x.$$

This is the component-wise statement of a matrix equation

$$A\alpha = \varphi, \tag{12.3}$$

to determine the coefficients $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_n)^{\mathrm{T}}$, where $A = \{a_{i,k}, i, k = 0, 1, \dots, n\}$, $\varphi = (f_0, f_1, \dots, f_n)^{\mathrm{T}}$,

$$a_{i,k} = \int_a^b w(x) x^{k+i} \, \mathrm{d}x \ \text{ and } \ f_i = \int_a^b w(x) f(x) x^i \, \mathrm{d}x.$$

The system (12.3) is called the set of **normal equations**.

Example: the best least-squares approximation to e^x on $[0, 1]$ from Π_1 in $\mathrm{L}^2(0, 1)$, with inner product $\langle f, g\rangle = \displaystyle\int_a^b f(x)g(x)\,\mathrm{d}x$.

Example: the best least-squares approximation to e^x on $[0,1]$ from Π_1 in $\mathrm{L}^2(0,1)$, with inner product $\langle f, g\rangle = \int_a^b f(x)g(x)\,\mathrm{d}x$. We want

$$\int_0^1 [\mathrm{e}^x - (\alpha_0 1 + \alpha_1 x)]1\,\mathrm{d}x = 0 \quad \text{and} \quad \int_0^1 [\mathrm{e}^x - (\alpha_0 1 + \alpha_1 x)]x\,\mathrm{d}x = 0.$$

Example: the best least-squares approximation to e^x on $[0,1]$ from Π_1 in $\mathrm{L}^2(0,1)$, with inner product $\langle f, g \rangle = \int_a^b f(x)g(x)\,\mathrm{d}x$. We want

$$\int_0^1 [\mathrm{e}^x - (\alpha_0 1 + \alpha_1 x)] 1\,\mathrm{d}x = 0 \text{ and } \int_0^1 [\mathrm{e}^x - (\alpha_0 1 + \alpha_1 x)] x\,\mathrm{d}x = 0.$$

$\iff$

$$\begin{aligned} \alpha_0 \int_0^1 \mathrm{d}x + \alpha_1 \int_0^1 x\,\mathrm{d}x &= \int_0^1 \mathrm{e}^x\,\mathrm{d}x \\ \alpha_0 \int_0^1 x\,\mathrm{d}x + \alpha_1 \int_0^1 x^2\,\mathrm{d}x &= \int_0^1 \mathrm{e}^x x\,\mathrm{d}x \end{aligned}$$

Example: the best least-squares approximation to e^x on $[0,1]$ from Π_1 in $\mathrm{L}^2(0,1)$, with inner product $\langle f, g\rangle = \int_a^b f(x)g(x)\,\mathrm{d}x$. We want

$$\int_0^1 [\mathrm{e}^x - (\alpha_0 1 + \alpha_1 x)]1\,\mathrm{d}x = 0 \text{ and } \int_0^1 [\mathrm{e}^x - (\alpha_0 1 + \alpha_1 x)]x\,\mathrm{d}x = 0.$$

$\Longleftrightarrow$

$$\alpha_0 \int_0^1 \mathrm{d}x + \alpha_1 \int_0^1 x\,\mathrm{d}x = \int_0^1 \mathrm{e}^x\,\mathrm{d}x$$

$$\alpha_0 \int_0^1 x\,\mathrm{d}x + \alpha_1 \int_0^1 x^2\,\mathrm{d}x = \int_0^1 \mathrm{e}^x x\,\mathrm{d}x$$

i.e., $$\begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix} = \begin{pmatrix} \mathrm{e} - 1 \\ 1 \end{pmatrix}$$

Example: the best least-squares approximation to e^x on $[0,1]$ from Π_1 in $\mathrm{L}^2(0,1)$, with inner product $\langle f,g\rangle = \int_a^b f(x)g(x)\,\mathrm{d}x$. We want

$$\int_0^1 [\mathrm{e}^x - (\alpha_0 1 + \alpha_1 x)]1\,\mathrm{d}x = 0 \quad \text{and} \quad \int_0^1 [\mathrm{e}^x - (\alpha_0 1 + \alpha_1 x)]x\,\mathrm{d}x = 0.$$

$\iff$

$$\begin{aligned} \alpha_0 \int_0^1 \mathrm{d}x + \alpha_1 \int_0^1 x\,\mathrm{d}x &= \int_0^1 \mathrm{e}^x\,\mathrm{d}x \\ \alpha_0 \int_0^1 x\,\mathrm{d}x + \alpha_1 \int_0^1 x^2\,\mathrm{d}x &= \int_0^1 \mathrm{e}^x x\,\mathrm{d}x \end{aligned}$$

i.e., $$\begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix} = \begin{pmatrix} \mathrm{e}-1 \\ 1 \end{pmatrix} \implies \begin{cases} \alpha_0 = 4\mathrm{e} - 10 \\ \alpha_1 = 18 - 6\mathrm{e} \end{cases}$$

Example: the best least-squares approximation to e^x on $[0,1]$ from Π_1 in $\mathrm{L}^2(0,1)$, with inner product $\langle f, g\rangle = \int_a^b f(x)g(x)\,\mathrm{d}x$. We want

$$\int_0^1 [\mathrm{e}^x - (\alpha_0 1 + \alpha_1 x)]1\,\mathrm{d}x = 0 \text{ and } \int_0^1 [\mathrm{e}^x - (\alpha_0 1 + \alpha_1 x)]x\,\mathrm{d}x = 0.$$

$\Longleftrightarrow$

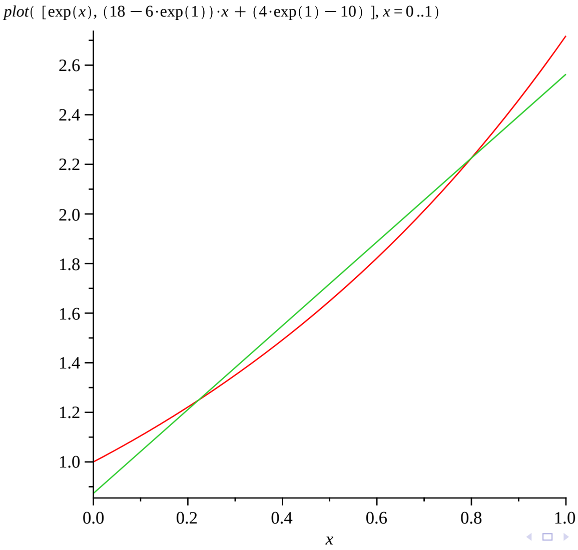
$$\alpha_0 \int_0^1 \mathrm{d}x + \alpha_1 \int_0^1 x\,\mathrm{d}x = \int_0^1 \mathrm{e}^x\,\mathrm{d}x$$

$$\alpha_0 \int_0^1 x\,\mathrm{d}x + \alpha_1 \int_0^1 x^2\,\mathrm{d}x = \int_0^1 \mathrm{e}^x x\,\mathrm{d}x$$

i.e., $$\begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix} = \begin{pmatrix} \mathrm{e}-1 \\ 1 \end{pmatrix} \Longrightarrow \begin{cases} \alpha_0 = 4\mathrm{e} - 10 \\ \alpha_1 = 18 - 6\mathrm{e} \end{cases}$$

so $p_1(x) := (18 - 6\mathrm{e})x + (4\mathrm{e} - 10)$ is the best approximation.

We plot $f(x) = e^x$ and $p_1(x)$ for $x \in [0, 1]$ using Maple



Proof that the coefficient matrix A is nonsingular will now establish existence and uniqueness of (weighted) $\|\cdot\|_2$ best-approximation.

Theorem

The coefficient matrix A is nonsingular.

Proof. Suppose not

Proof. Suppose not $\implies \exists \alpha \neq 0$ with $A\alpha = 0$

Proof. Suppose not $\implies \exists\alpha \neq 0$ with $A\alpha = 0 \implies \alpha^{\mathrm{T}} A\alpha = 0$

Proof. Suppose not $\implies \exists \alpha \neq 0$ with $A\alpha = 0 \implies \alpha^{\mathrm{T}} A\alpha = 0$

$$\iff \quad \sum_{i=0}^{n} \alpha_i (A\alpha)_i = 0$$

Proof. Suppose not $\Longrightarrow$ $\exists \alpha \neq 0$ with $A\alpha = 0 \Longrightarrow \alpha^{\mathrm{T}} A\alpha = 0$

$$\iff \quad \sum_{i=0}^{n} \alpha_i (A\alpha)_i = 0 \quad \iff \quad \sum_{i=0}^{n} \alpha_i \sum_{k=0}^{n} a_{ik}\alpha_k = 0,$$

Proof. Suppose not $\implies \exists \alpha \neq 0$ with $A\alpha = 0 \implies \alpha^{\mathrm{T}} A\alpha = 0$

$$\iff \sum_{i=0}^{n} \alpha_i (A\alpha)_i = 0 \iff \sum_{i=0}^{n} \alpha_i \sum_{k=0}^{n} a_{ik}\alpha_k = 0,$$

and using the definition $a_{ik} = \displaystyle\int_a^b w(x) x^k x^i \,\mathrm{d}x$,

Proof. Suppose not $\implies \exists \alpha \neq 0$ with $A\alpha = 0 \implies \alpha^{\mathrm{T}} A\alpha = 0$

$$\iff \sum_{i=0}^{n} \alpha_i (A\alpha)_i = 0 \iff \sum_{i=0}^{n} \alpha_i \sum_{k=0}^{n} a_{ik}\alpha_k = 0,$$

and using the definition $a_{ik} = \displaystyle\int_a^b w(x) x^k x^i \, \mathrm{d}x$,

$$\iff \sum_{i=0}^{n} \alpha_i \sum_{k=0}^{n} \left(\int_a^b w(x) x^k x^i \, \mathrm{d}x \right) \alpha_k = 0.$$

Rearranging gives

Rearranging gives

$$\int_a^b w(x) \left(\sum_{i=0}^{n} \alpha_i x^i \right) \left(\sum_{k=0}^{n} \alpha_k x^k \right) \mathrm{d}x = 0$$

Rearranging gives

$$\int_a^b w(x) \left(\sum_{i=0}^{n} \alpha_i x^i \right) \left(\sum_{k=0}^{n} \alpha_k x^k \right) \mathrm{d}x = 0 \text{ or } \int_a^b w(x) \left(\sum_{i=0}^{n} \alpha_i x^i \right)^2 \mathrm{d}x = 0$$

Rearranging gives

$$\int_a^b w(x)\left(\sum_{i=0}^n \alpha_i x^i\right)\left(\sum_{k=0}^n \alpha_k x^k\right) \mathrm{d}x = 0 \text{ or } \int_a^b w(x)\left(\sum_{i=0}^n \alpha_i x^i\right)^2 \mathrm{d}x = 0$$

which implies that $\displaystyle\sum_{i=0}^n \alpha_i x^i \equiv 0$

Rearranging gives

$$\int_a^b w(x) \left(\sum_{i=0}^{n} \alpha_i x^i \right) \left(\sum_{k=0}^{n} \alpha_k x^k \right) \mathrm{d}x = 0 \;\text{ or }\; \int_a^b w(x) \left(\sum_{i=0}^{n} \alpha_i x^i \right)^2 \mathrm{d}x = 0$$

which implies that $\sum_{i=0}^{n} \alpha_i x^i \equiv 0$ and thus $\alpha_i = 0$ for $i = 0, 1, \ldots, n$.

Rearranging gives

$$\int_a^b w(x) \left(\sum_{i=0}^{n} \alpha_i x^i \right) \left(\sum_{k=0}^{n} \alpha_k x^k \right) \mathrm{d}x = 0 \quad \text{or} \quad \int_a^b w(x) \left(\sum_{i=0}^{n} \alpha_i x^i \right)^2 \mathrm{d}x = 0$$

which implies that $\sum_{i=0}^{n} \alpha_i x^i \equiv 0$ and thus $\alpha_i = 0$ for $i = 0, 1, \ldots, n$.

This contradicts the initial supposition, and thus A is nonsingular. □

Orthogonal Polynomials

Gram–Schmidt orthogonalization procedure:

Orthogonal Polynomials

Gram–Schmidt orthogonalization procedure: the solution of the normal equations $A\alpha = \varphi$ for best least-squares polynomial approximation would be easy if A were diagonal.

Orthogonal Polynomials

Gram–Schmidt orthogonalization procedure: the solution of the normal equations $A\alpha = \varphi$ for best least-squares polynomial approximation would be easy if A were diagonal.

Idea: Instead of

$$\{1, x, x^2, \ldots, x^n\}$$

as a basis for Π_n, suppose we have a basis

$$\{\phi_0, \phi_1, \ldots, \phi_n\}.$$

Orthogonal Polynomials

Gram–Schmidt orthogonalization procedure: the solution of the normal equations $A\alpha = \varphi$ for best least-squares polynomial approximation would be easy if A were diagonal.

Idea: Instead of

$$\{1, x, x^2, \ldots, x^n\}$$

as a basis for Π_n, suppose we have a basis

$$\{\phi_0, \phi_1, \ldots, \phi_n\}.$$

We shall choose the basis $\{\phi_0, \phi_1, \ldots, \phi_n\}$ so as to ensure that the matrix A in the normal equations is diagonal.

With this choice of basis, we seek the polynomial of best approximation $p_n(x) = \sum_{k=0}^{n} \beta_k \phi_k(x)$. The normal equations then become

$$\int_a^b w(x) \left(f(x) - \sum_{k=0}^{n} \beta_k \phi_k(x) \right) \phi_i(x) \, \mathrm{d}x = 0, \quad i = 0, 1, \ldots, n,$$

With this choice of basis, we seek the polynomial of best approximation $p_n(x) = \sum_{k=0}^{n} \beta_k \phi_k(x)$. The normal equations then become

$$\int_a^b w(x) \left(f(x) - \sum_{k=0}^{n} \beta_k \phi_k(x) \right) \phi_i(x) \, \mathrm{d}x = 0, \quad i = 0, 1, \ldots, n,$$

or equivalently

$$\sum_{k=0}^{n} \left(\int_a^b w(x) \phi_k(x) \phi_i(x) \, \mathrm{d}x \right) \beta_k = \int_a^b w(x) f(x) \phi_i(x) \mathrm{d}x, \quad i = 0, 1, \ldots, n,$$

With this choice of basis, we seek the polynomial of best approximation $p_n(x) = \sum_{k=0}^{n} \beta_k \phi_k(x)$. The normal equations then become

$$\int_a^b w(x) \left(f(x) - \sum_{k=0}^{n} \beta_k \phi_k(x) \right) \phi_i(x) \, \mathrm{d}x = 0, \quad i = 0, 1, \ldots, n,$$

or equivalently

$$\sum_{k=0}^{n} \left(\int_a^b w(x) \phi_k(x) \phi_i(x) \, \mathrm{d}x \right) \beta_k = \int_a^b w(x) f(x) \phi_i(x) \mathrm{d}x, \quad i = 0, 1, \ldots, n,$$

i.e.,

$$A\beta = \varphi, \tag{12.4}$$

where $\beta = (\beta_0, \beta_1, \ldots, \beta_n)^{\mathrm{T}}$, $\varphi = (f_1, f_2, \ldots, f_n)^{\mathrm{T}}$ and now

$$a_{i,k} = \int_a^b w(x) \phi_k(x) \phi_i(x) \, \mathrm{d}x \qquad \text{and} \qquad f_i = \int_a^b w(x) f(x) \phi_i(x) \, \mathrm{d}x.$$

So A is diagonal if

$$\langle \phi_i, \phi_k \rangle = \int_a^b w(x)\phi_i(x)\phi_k(x)\,\mathrm{d}x \quad \begin{cases} = 0 & i \neq k \text{ and} \\ \neq 0 & i = k. \end{cases}$$

So A is diagonal if

$$\langle \phi_i, \phi_k \rangle = \int_a^b w(x)\phi_i(x)\phi_k(x)\,\mathrm{d}x \quad \begin{cases} = 0 & i \neq k \text{ and} \\ \neq 0 & i = k. \end{cases}$$

We can create such a set of **orthogonal polynomials**

$$\{\phi_0, \phi_1, \dots, \phi_n, \dots\},$$

with $\phi_i \in \Pi_i$ for each i, by the Gram–Schmidt procedure, which is based on the following lemma.

Lemma

Suppose that $\{\phi_0, \phi_1, \ldots, \phi_k\}$*, with* $\phi_i \in \Pi_i$ *for each* i*, are orthogonal with respect to the inner product* $\langle f, g\rangle = \int_a^b w(x) f(x) g(x)\, \mathrm{d}x$*. Then,*

$$\phi_{k+1}(x) = x^{k+1} - \sum_{i=0}^{k} \lambda_i \phi_i(x)$$

satisfies

$$\langle \phi_{k+1}, \phi_j\rangle = \int_a^b w(x) \phi_{k+1}(x) \phi_j(x)\, \mathrm{d}x = 0, \qquad j = 0, 1, \ldots, k,$$

when

$$\lambda_j = \frac{\langle x^{k+1}, \phi_j\rangle}{\langle \phi_j, \phi_j\rangle}, \qquad j = 0, 1, \ldots, k.$$

Proof. For any j, $0 \le j \le k$,

$$\begin{aligned}\langle \phi_{k+1}, \phi_j \rangle = & \langle x^{k+1}, \phi_j \rangle - \sum_{i=0}^{k} \lambda_i \langle \phi_i, \phi_j \rangle \\ = & \langle x^{k+1}, \phi_j \rangle - \lambda_j \langle \phi_j, \phi_j \rangle\end{aligned}$$

by the orthogonality of ϕ_i and ϕ_j, $i \neq j$,

$$= 0$$ by definition of λ_j. □

Notes:

1. The Gram–Schmidt procedure does the above for $k = 0, 1, \ldots, n$ successively.
2. ϕ_k is always of exact degree k, so $\{\phi_0, \phi_1, \ldots, \phi_\ell\}$ is a basis for Π_ℓ for every $\ell \geq 0$.
3. ϕ_k can be normalised/scaled to satisfy $\langle \phi_k, \phi_k \rangle = 1$ or to be monic.