Dr A. Iserles
Easter 1997

Numerical Analysis – Lecture 10

7.3 The three-term recurrence relation

How to construct orthogonal polynomials? (7.1) might help, but it suffers from the same problem as the Gram–Schmidt algorithm in Euclidean spaces: loss of accuracy due to ill-conditioning. A considerably better procedure follows from our next theorem.

Theorem Monic orthogonal polynomials are given by the formula

$$\begin{aligned} &p_{-1}(x) \equiv 0, \qquad p_0(x) \equiv 1, \\ &p_{n+1}(x) = (x-\alpha_n)p_n(x) - \beta_n p_{n-1}(x), \qquad n = 0,1,\ldots, \end{aligned} \tag{7.1}$$

where

$$\alpha_n := \frac{\langle p_n, xp_n\rangle}{\langle p_n, p_n\rangle}, \qquad \beta_n = \frac{\langle p_n, p_n\rangle}{\langle p_{n-1}, p_{n-1}\rangle} > 0.$$

Proof. Pick $n \geq 0$ and let

$$\psi(x) := p_{n+1}(x) - (x-\alpha_n)p_n(x) + \beta_n p_{n-1}(x).$$

Since p_n and p_{n+1} are monic, it follows that $\psi \in \mathbb{P}_n[x]$. Moreover,

$$\langle \psi, p_\ell\rangle = \langle p_{n+1}, p_\ell\rangle - \langle p_n, (x-\alpha_n)p_\ell\rangle + \beta_n\langle p_{n-1}, p_\ell\rangle = 0, \qquad \ell = 0,1,\ldots,n-2.$$

Because of monicity, $xp_{n-1} = p_n + q$, where $q \in \mathbb{P}_{n-1}[x]$. Thus, from the definition of α_n, β_n,

$$\begin{aligned} \langle \psi, p_{n-1}\rangle &= -\langle p_n, xp_{n-1}\rangle + \beta_n\langle p_{n-1}, p_{n-1}\rangle = -\langle p_n, p_n\rangle + \beta_n\langle p_{n-1}, p_{n-1}\rangle = 0, \\ \langle \psi, p_n\rangle &= -\langle xp_n, p_n\rangle + \alpha_n\langle p_n, p_n\rangle = 0. \end{aligned}$$

Every $p \in \mathbb{P}_n[x]$ that obeys $\langle p, p_\ell\rangle = 0$, $\ell = 0,1,\ldots,n$, must necessarily be the zero polynomial. For suppose that it is not so and let x^s be the highest coefficient of x in p. Then $\langle p, p_s\rangle \neq 0$. We deduce that $\psi \equiv 0$, hence (7.1) is true. □

Example *Chebyshev polynomials* We choose the scalar product

$$\langle f, g\rangle := \int_{-1}^{1} f(x)g(x)\frac{\mathrm{d}x}{\sqrt{1-x^2}}, \qquad f, g \in C[-1,1]$$

and define $T_n \in \mathbb{P}_n[x]$ *via* the relation $T_n(\cos\theta) = \cos n\theta$ – hence $T_0(x) \equiv 1$, $T_1(x) = x$, $T_2(x) = 2x^2 - 1$ etc. Changing the integration variable,

$$\begin{aligned} \langle T_n, T_m\rangle &= \int_{-1}^{1} T_n(x)T_m(x)\frac{\mathrm{d}x}{\sqrt{1-x^2}} = \int_0^\pi \cos n\theta \cos m\theta \,\mathrm{d}\theta \\ &= \tfrac{1}{2}\int_0^\pi \{\cos(n+m)\theta + \cos(n-m)\theta\}\,\mathrm{d}\theta = 0 \end{aligned}$$

whenever $n \neq m$. The recurrence relation for Chebyshev polynomials is

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x),$$

as can be verified at once from the identity

$$\cos(n+1)\theta + \cos(n-1)\theta = 2\cos\theta\cos n\theta.$$

Note that the T_ns aren't monic, hence the inconsistency with (7.1).

7.4 Least-squares polynomial fitting

Given $f \in C[a,b]$ and a scalar product $\langle g,h\rangle = \int_a^b w(x)g(x)h(x)\,\mathrm{d}x$, we wish to pick $p \in \mathbb{P}_n[x]$ so as to minimize $\langle f-p, f-p\rangle$. Again, $w(x) > 0$ for $x \in (a,b)$. Intuitively speaking, p approximates f and is an alternative to an interpolating polynomial. (The situation is similar to the one that we have already encountered in numerical linear algebra.)
Let $p_0, p_1, \ldots, p_n$ be orthogonal polynomials w.r.t. the underlying inner product, $p_\ell \in \mathbb{P}_\ell[x]$. We can represent $p = \sum_{k=0}^n c_k p_k$ for some $c_0, c_1, \ldots, c_n \in \mathbb{R}$, hence, by orthogonality,

$$\langle f-p, f-p\rangle = \left\langle f - \sum_{k=0}^n c_k p_k, f - \sum_{k=0}^n c_k p_k \right\rangle = \langle f,f\rangle - 2\sum_{k=0}^n c_k \langle p_k, f\rangle + \sum_{k=0}^n c_k^2 \langle p_k, p_k\rangle.$$

To derive optimal $c_0, c_1, \ldots, c_n$ we seek to minimize the last expression. Since

$$\tfrac{1}{2}\frac{\partial}{\partial c_k}\langle f-p, f-p\rangle = -\langle p_k, f\rangle + c_k\langle p_k, p_k\rangle, \qquad k = 0,1,\ldots,n,$$

setting the gradient to zero yields

$$p(x) = \sum_{k=0}^n \frac{\langle p_k, f\rangle}{\langle p_k, p_k\rangle} p_k(x). \tag{7.2}$$

Note that

$$\langle f-p, f-p\rangle = \langle f,f\rangle - \sum_{k=0}^n \{2c_k\langle p_k, f\rangle - c_k^2\langle p_k, p_k\rangle\} = \langle f,f\rangle - \sum_{k=0}^n \frac{\langle p_k, f\rangle^2}{\langle p_k, p_k\rangle}. \tag{7.3}$$

How to choose n? Note that $c_k = \langle p_k, f\rangle / \langle p_k, p_k\rangle$ is independent of n. Thus, we can continue to add terms to (7.2) until $\langle f-p, f-p\rangle$ is below specified *tolerance* ε. Because of (7.3), we need to pick n so that $\langle f,f\rangle - \varepsilon < \sum_{k=0}^n \langle p_k, f\rangle^2 / \langle p_k, p_k\rangle$.

Theorem *(The Parseval identity)* Let $[a,b]$ be finite. Then

$$\sum_{k=0}^\infty \frac{\langle p_k, f\rangle^2}{\langle p_k, p_k\rangle} = \langle f,f\rangle. \tag{7.4}$$

Incomplete proof. Let $\sigma_n := \sum_{k=0}^n \langle p_k, f\rangle^2 / \langle p_k, p_k\rangle$, $n = 0,1,\ldots$, hence $\langle f-p, f-p\rangle = \langle f,f\rangle - \sigma_n \ge 0$. The sequence $\{\sigma\}_{n=0}^\infty$ increases monotonically and $\sigma_n \le \langle f,f\rangle$ implies that $\lim_{n\to\infty}\sigma_n$ exists. According to the *Weierstrass theorem*, any function in $C[a,b]$ can be approximated arbitrarily close by a polynomial, hence $\lim_{n\to\infty}\langle f-p, f-p\rangle = 0$ and we deduce that $\sigma_n \stackrel{n\to\infty}{\longrightarrow} \langle f,f\rangle$ and (7.4) is true. □