Lecture 16: From Lanczos to Gauss quadrature

School of Mathematical Sciences, Xiamen University

1. Orthogonal polynomials

- Replace $\mathbb{C}^n$ by $L^2[-1,1]$, a vector space of real-valued functions on $[-1,1]$. The inner product of two functions $u, v \in L^2[-1,1]$ is defined by

$$\langle u, v\rangle = \int_{-1}^{1} u(x)v(x)\mathrm{d}x,$$

and the norm of a function $u \in L^2[-1,1]$ is $\|u\| = \langle u, u\rangle^{1/2}$.

Proposition 1

The linear operator $\mathbf{A} : L^2[-1,1] \to L^2[-1,1]$ defined by

$$(\mathbf{A}u)(x) = xu(x)$$

is self-adjoint with respect to the given inner product.

Proof. Note that

$$\langle \mathbf{A}u, v\rangle = \int_{-1}^{1} (\mathbf{A}u)(x)v(x)\mathrm{d}x = \int_{-1}^{1} u(x)(\mathbf{A}v)(x)\mathrm{d}x = \langle u, \mathbf{A}v\rangle. \quad \square$$

- The Lanczos process ($\mathbf{r} = 1$ and $\mathbf{A} = x$) becomes the procedure for constructing orthogonal polynomials via a three-term recurrence relation: $x\begin{bmatrix} q_1(x) & \cdots & q_j(x) \end{bmatrix} = \begin{bmatrix} q_1(x) & \cdots & q_{j+1}(x) \end{bmatrix} \widetilde{\mathbf{T}}_j$.

Algorithm: Lanczos process for orthogonal polynomials

$\beta_0 = 0,\ q_0(x) = 0,\ q_1(x) = 1/\sqrt{2}$

for $j = 1, 2, 3, \ldots,$

  $v(x) = xq_j(x)$

  $v(x) = v(x) - \beta_{j-1}q_{j-1}(x)$

  $\alpha_j = \langle v, q_j \rangle$

  $v(x) = v(x) - \alpha_j q_j(x)$

  $\beta_j = \|v\|$

  $q_{j+1}(x) = v(x)/\beta_j$

end

$$\widetilde{\mathbf{T}}_j = \begin{bmatrix} \alpha_1 & \beta_1 & & \\ \beta_1 & \alpha_2 & \ddots & \\ & \ddots & \ddots & \beta_{j-1} \\ & & \beta_{j-1} & \alpha_j \\ & & & \alpha_{j+1} \end{bmatrix}$$

Remark 2

We have $\langle q_i, q_j \rangle = \displaystyle\int_{-1}^{1} q_i(x)q_j(x)\mathrm{d}x = \delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$

Remark 3

The function $q_j(x)$ *is a scalar multiple of the Legendre polynomial* $P_j(x)$ *of degree* $j-1$ *(note that* $P_j(1)=1$*), i.e.,*

$$q_j(x) = q_j(1)P_j(x).$$

Remark 4

The three-term recurrence takes the form

$$xq_j(x) = \beta_{j-1}q_{j-1}(x) + \alpha_j q_j(x) + \beta_j q_{j+1}(x).$$

The entries $\{\alpha_j\}$ *and* $\{\beta_j\}$ *are known analytically:*

$$\alpha_j = 0, \qquad \beta_j = \frac{1}{2}(1-(2j)^{-2})^{-1/2}.$$

- The tridiagonal matrices $\{\mathbf{T}_j\}$ in Lanczos process are known as *Jacobi matrices* in the context of orthogonal polynomials.

1.1. Comparison to Gram–Schmidt

Algorithm: Gram–Schmidt for orthogonal polynomials

```
for j = 1, 2, 3, ...
    q_j(x) = x^{j-1}
    for i = 1 to j-1
        r_ij = <x^{j-1}, q_i>
        q_j(x) = q_j(x) - r_ij q_i(x)
    end
    r_jj = ||q_j||
    q_j(x) = q_j(x)/r_jj
end
```

for $j = 1, 2, 3, \cdots$

  $q_j(x) = x^{j-1}$

  for $i = 1$ **to** $j-1$

    $r_{ij} = \langle x^{j-1}, q_i \rangle$

    $q_j(x) = q_j(x) - r_{ij} q_i(x)$

  end

  $r_{jj} = \|q_j\|$

  $q_j(x) = q_j(x)/r_{jj}$

end

Remark 5

The above algorithm constructs the continuous QR *factorizations of the "Krylov matrix"*

$$\mathbf{K}_\infty = \begin{bmatrix} 1 & x & x^2 & x^3 & \cdots \end{bmatrix},$$

which is obtained by setting $\mathbf{r} = 1$ *and* $\mathbf{A} = x$.

Remark 6

Lanczos process and Gram–Schmidt both obtain the same sequence of orthogonal polynomials $\{q_j\}$.

Remark 7

If the inner product is modified by the inclusion of a nonconstant positive weight function $w(x)$ *in the integrand, then one obtains other families of orthogonal polynomials such as Chebyshev polynomials and Jacobi polynomials.*

2. Orthogonal polynomials approximation problem

- Find a monic polynomial p^j of degree j such that

$$\|p^j(x)\| = \min_{\text{monic } p, \ \deg(p)=j} \|p(x)\|.$$

According to Theorem 6 of Lecture 15, the unique solution is the characteristic polynomial of the matrix $\mathbf{T}_j$.

Theorem 8

Let $p^j(x)$ be the characteristic polynomial of $\mathbf{T}_j$. Then for $j = 0, 1, \cdots$,

$$p^j(x) = \rho_j q_{j+1}(x),$$

where ρ_j is a constant.

Proof. Any monic $p(x)$ of degree j can be written as

$$p(x) = \rho_j q_{j+1}(x) + \sum_{i=1}^{j} y_i q_i(x),$$

where ρ_j is a constant – the inverse of the leading coefficient of $q_{j+1}(x)$. Due to

$$\|p(x)\| = (\rho_j^2 + \|\mathbf{y}\|_2^2)^{1/2},$$

the minimum is obtained by setting $\mathbf{y} = \mathbf{0}$. $\square$

Corollary 9

The zeros of $q_{j+1}(x)$ are the eigenvalues of $\mathbf{T}_j$. These j zeros are distinct and lie in the open interval $(-1, 1)$.

Proof. All eigenvalues of $\mathbf{T}_j$ are distinct. Assume that $k < j$. For any $\{x_i\}_{i=1}^k$, we have

$$\int_{-1}^{1} q_{j+1}(x)\mathrm{d}x = 0, \quad \int_{-1}^{1} q_{j+1}(x)\prod_{i=1}^{k}(x - x_i)\mathrm{d}x = 0.$$

The first equality shows that there exists at least one zero in $(-1, 1)$. Now assume there are only $k < j$ distinct zeros in $(-1, 1)$, denoted by $\{x_i\}_{i=1}^k$. Consider the polynomial $q_{j+1}(x)\prod_{i=1}^{k}(x - x_i)$, which has constant sign in $(-1, 1)$. This is a contradiction of the second equality. □

3. Gauss–Legendre quadrature

- Numerical quadrature: consider a j-point quadrature formula

$$\mathcal{I}_j(f) = \sum_{i=1}^{j} w_i f(x_i) \quad \text{for} \quad \mathcal{I}(f) = \int_{-1}^{1} f(x)\mathrm{d}x.$$

Theorem 10

Let the nodes $\{x_i\}_{i=1}^{j}$ be an arbitrary set of j distinct points in $[-1,1]$. Then there is a unique choice of weights $\{w_i\}_{i=1}^{j}$ with the property that the quadrature formula ($\mathcal{I}_j(f) = \sum_{i=1}^{j} w_i f(x_i)$) has order of accuracy at least $j-1$ in the sense that it is exact (i.e., $\mathcal{I}_j(f) = \int_{-1}^{1} f(x)\mathrm{d}x$) if $f(x)$ is any polynomial of degree $\leq j-1$. The weights $\{w_i\}_{i=1}^{j}$ are given by

$$w_i = \int_{-1}^{1} \ell_i(x)\mathrm{d}x, \quad \ell_i(x) = \prod_{k=1,k\neq i}^{j} (x - x_k) \Big/ \prod_{k=1,k\neq i}^{j} (x_i - x_k).$$

- Gauss–Legendre quadrature: $\{x_i\}_{i=1}^{j}$ are the zeros of $q_{j+1}(x)$.

Theorem 11

The j-point Gauss–Legendre quadrature formula has order of accuracy exactly $2j-1$, and no quadrature formula has order of accuracy higher than this.

Proof. Consider the polynomial

$$f(x) = \prod_{i=1}^{j}(x - x_i)^2, \qquad \mathcal{I}(f) = \int_{-1}^{1} f(x)\mathrm{d}x > 0.$$

Note that $\mathcal{I}_j(f) = 0$ since $f(x_i) = 0$. Thus the quadrature formula has order of accuracy $\leq 2j-1$. For any $f(x) \in \mathbb{P}_{2j-1}$, it can be factored in the form

$$f(x) = g(x)q_{j+1}(x) + r(x),$$

where $g(x) \in \mathbb{P}_{j-1}$ and $r(x) \in \mathbb{P}_{j-1}$. (In fact, $r(x)$ is the unique polynomial interpolant to $f(x)$ in the points $\{x_i\}$.)

Since $q_{j+1}(x)$ is orthogonal to all polynomials of lower degree, we have

$$\mathcal{I}(gq_{j+1}) = 0.$$

At the same time, since

$$g(x_i)q_{j+1}(x_i) = 0$$

for each x_i, we have

$$\mathcal{I}_j(gq_{j+1}) = 0.$$

Since $\mathcal{I}$ and $\mathcal{I}_j$ are linear operators, these identities imply

$$\mathcal{I}(f) = \mathcal{I}(r) \quad \text{and} \quad \mathcal{I}_j(f) = \mathcal{I}_j(r).$$

Therefore, by Theorem 10, i.e.,

$$\mathcal{I}(r) = \mathcal{I}_j(r),$$

we have

$$\mathcal{I}(f) = \mathcal{I}_j(f). \quad \square$$

Theorem 12

Let $\mathbf{T}_j$ be the $j \times j$ Jacobi matrix. Let $\mathbf{T}_j = \mathbf{V}\mathbf{D}\mathbf{V}^\top$ be an orthogonal diagonalization of $\mathbf{T}_j$ with

$$\mathbf{D} = \operatorname{diag}\{\lambda_1, \cdots, \lambda_j\}, \quad \mathbf{V} = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_j \end{bmatrix}.$$

Then the nodes and weights of the Gauss–Legendre quadrature formula are given by

$$x_i = \lambda_i, \quad w_i = 2(\mathbf{v}_i)_1^2, \quad i = 1, \cdots, j.$$

- G. H. Golub and J. H. Welsch

  Calculation of Gauss quadrature rules, Math. Comp. 23 (1969).

  The famous $O(j^2)$ algorithm for Gauss quadrature nodes and weights via a tridiagonal Jacobi matrix eigenvalue problem.
- G. H. Golub and G. Meurant

  Matrices, Moments and Quadrature with Applications

  Princeton University Press, 2010