UPPSALA UNIVERSITET

Error Analysis of Gauss-Legendre-Lobatto Quadrature and the Conjugate Gradient Method

David Meadon - David Niemelä - Zhenlu Sun - Usama Zafar

Gauss-Legendre-Lobatto Quadrature

Introduction

Numerical integration tries to solve: Find the integral,

$$\mathfrak{I}(f) = \int_a^b f(x)\mathrm{d}x$$

for some function $f(x) \in \mathrm{C}^0([a, b])$.

Solution: Approximate the integral with

$$\mathrm{I}_n(f) = \sum_{i=0}^{n} \bar{\alpha}_i f(\bar{x}_i)$$

with some weights and nodes, $\bar{\alpha}_i$ and $\bar{x}_i$.

Gauss-Legendre-Lobatto Quadrature

Introduction

Some basic methods:

- Midpoint rule:
  $\bar{\alpha}_i = x_i - x_{i-1}, f(\bar{x}_i) = f\left(\frac{x_i - x_{i-1}}{2}\right)$

- Trapezoidal rule:
  $\bar{\alpha}_i = x_i - x_{i-1}, f(\bar{x}_i) = \frac{f(x_i) + f(x_{i-1})}{2}$

Gauss-Legendre-Lobatto Quadrature

Introduction

Some basic methods:

- Midpoint rule:
  $\bar{\alpha}_i = x_i - x_{i-1}, f(\bar{x}_i) = f\left(\frac{x_i - x_{i-1}}{2}\right)$ - Piecewise Constant Interpolation
- Trapezoidal rule:
  $\bar{\alpha}_i = x_i - x_{i-1}, f(\bar{x}_i) = \frac{f(x_i) + f(x_{i-1})}{2}$ - Piecewise Linear Interpolation

Gauss-Legendre-Lobatto Quadrature

Introduction

Some basic methods:

- Midpoint rule:
  $\bar{\alpha}_i = x_i - x_{i-1}, f(\bar{x}_i) = f\left(\frac{x_i - x_{i-1}}{2}\right)$ - Piecewise Constant Interpolation
- Trapezoidal rule:
  $\bar{\alpha}_i = x_i - x_{i-1}, f(\bar{x}_i) = \frac{f(x_i) + f(x_{i-1})}{2}$ - Piecewise Linear Interpolation

In this course we have seen we can do better than these interpolations: Orthogonal Polynomials!

Gauss-Legendre-Lobatto Quadrature

Legendre Polynomials

UPPSALA UNIVERSITET

The polynomials we will use:

$$L_k(x) = \frac{1}{2^k} \sum_{l=0}^{\lfloor k/2 \rfloor} (-1)^l \binom{k}{l} \binom{2k-2l}{k} x^{k-2l},$$

Orthogonal polynomials over the interval $(-1, 1)$. We have that

$$\left(L_i, L_j\right)_{\mathrm{L}^2(-1,1)} = \delta_{ij} \left(i + \frac{1}{2}\right)^{-1}$$

Gauss-Legendre-Lobatto Quadrature

Method

The Integral approximation of f over the interval $(-1, 1)$:

$$\int_{-1}^{1} f(x)\mathrm{d}x \approx \mathrm{I}_n^{GL}(f) = \sum_{i=0}^{n} \bar{\alpha}_i f(\bar{x}_i)$$

uses the nodes:

$\bar{x}_0 = -1$, $\bar{x}_n = 1$ and $\bar{x}_i$, $i = 1, \ldots, n-1$ are the roots of $L_n'(x)$,

and weights:

$$\bar{\alpha}_i = \frac{2}{n(n+1)} \frac{1}{[L_n(\bar{x}_i)]^2}$$

Gauss-Legendre-Lobatto Quadrature

Convergence Estimate

$$|\mathfrak{I}(f) - \mathrm{I}_n^{\mathrm{GL}}(f)| \leq Cn^{-s}\|f\|_s$$

where

$$\|f\|_s = \left(\sum_{k=0}^{s} \|f^{(k)}\|^2_{\mathrm{L}^2(-1,1)}\right)^{\frac{1}{2}}$$

That is the standard norm on the Sobolev space $\mathrm{H}^s(-1,1)$, the space of functions which have bounded L^2-norm on the functions and their weak derivatives up to order s.

Small note:

$$\int_a^b f(x)\mathrm{d}x = \frac{b-a}{2}\int_{-1}^{1} g(\xi)\mathrm{d}\xi, \quad g(\xi) = f(\frac{b-a}{2}\xi + \frac{a+b}{2})$$

Gauss-Legendre-Lobatto Quadrature

Convergence Estimate - proof

5.4 Legendre Approximations 281

For the *Gauss-Lobatto integration*, if (5.3.2) holds, then there exists a positive constant C independent of N such that

$$|(u,\phi)_w - (u,\phi)_N| \le C(\|u - P_{N-1}u\|_{L^2_w(-1,1)} + \|u - I_N u\|_{L^2_w(-1,1)})\|\phi\|_{L^2_w(-1,1)} \ . \tag{5.3.4b}$$

Actually we have

$$\begin{aligned}
&|(u,\phi)_w - (u,\phi)_N| \\
&\quad = |(u,\phi)_w - (P_{N-1}u,\phi)_w + (P_{N-1}u,\phi)_w - (I_N u,\phi)_N| \\
&\quad \le |(u - P_{N-1}u,\phi)_w| + |(P_{N-1}u - I_N u,\phi)_N| \quad \text{(by (2.2.25))} \\
&\quad \le C\left(\|u - P_{N-1}u\|_{L^2_w(-1,1)} + \|P_{N-1}u - I_N u\|_N\right)\|\phi\|_{L^2_w(-1,1)} \\
&\qquad\qquad \text{(by the Cauchy-Schwarz inequality and (5.3.2))} \\
&\quad \le C\left(2\|u - P_{N-1}u\|_{L^2_w(-1,1)} + \|u - I_N u\|_{L^2_w(-1,1)}\right)\|\phi\|_{L^2_w(-1,1)} \\
&\qquad\qquad \text{(by (5.3.2))}\ ;
\end{aligned}$$

whence, (5.3.4b) follows.

Figure: Excerpt from [1]

Gauss-Legendre-Lobatto Quadrature

Convergence Estimate - proof

UPPSALA UNIVERSITET

Remarks:

ϕ is some polynomial P_N, $w = 1$

Inner-product $(u, \phi)_N = \sum_{i=0}^{N} \alpha_i u(x_i) = (I_N u, \phi)_N$ (2.2.25) in [1]

exact: $(u, v)_N = (u, v)_w \quad uv \in P_{2N-1}$

Tools and "tricks" used:

Adding nothing, to get something we know more about.

Equivalence of norms, $C_1||\phi||_{L^2} \leq ||\phi||_N \leq C_2||\phi||_{L^2}$, (5.3.2) in [1]

Gauss-Legendre-Lobatto Quadrature

Convergence Estimate - proof

5.4 Legendre Approximations 281

For the *Gauss-Lobatto integration*, if (5.3.2) holds, then there exists a positive constant C independent of N such that

$$|(u,\phi)_w - (u,\phi)_N| \le C(\|u - P_{N-1}u\|_{L^2_w(-1,1)} + \|u - I_N u\|_{L^2_w(-1,1)})\|\phi\|_{L^2_w(-1,1)} \ . \qquad (5.3.4b)$$

Actually we have

$$\begin{aligned}
&|(u,\phi)_w - (u,\phi)_N| \\
&\quad = |(u,\phi)_w - (P_{N-1}u,\phi)_w + (P_{N-1}u,\phi)_w - (I_N u,\phi)_N| \\
&\quad \le |(u - P_{N-1}u,\phi)_w| + |(P_{N-1}u - I_N u,\phi)_N| \quad \text{(by (2.2.25))} \\
&\quad \le C\left(\|u - P_{N-1}u\|_{L^2_w(-1,1)} + \|P_{N-1}u - I_N u\|_N\right)\|\phi\|_{L^2_w(-1,1)} \\
&\qquad\qquad \text{(by the Cauchy-Schwarz inequality and (5.3.2))} \\
&\quad \le C\left(2\|u - P_{N-1}u\|_{L^2_w(-1,1)} + \|u - I_N u\|_{L^2_w(-1,1)}\right)\|\phi\|_{L^2_w(-1,1)} \\
&\qquad\qquad \text{(by (5.3.2))} \ ;
\end{aligned}$$

whence, (5.3.4b) follows.

Figure: Excerpt from [1]

Gauss-Legendre-Lobatto Quadrature

Convergence Estimate - proof

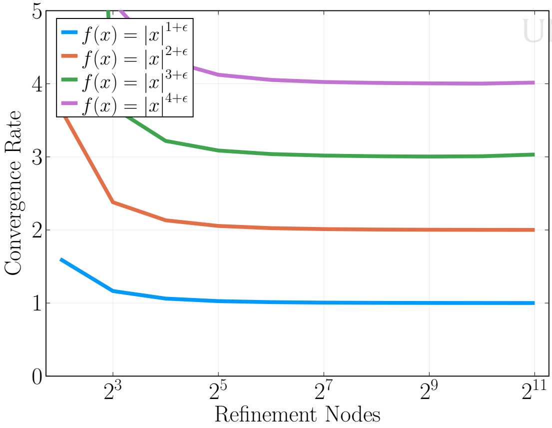


Figure: Convergence plot of differently smooth functions, $\epsilon = 10^{-5}$

Conjugate Gradient Method

Introduction

Conjugate Gradient Method is an iterative method for solving linear system of equation of the form:

$$Ax = b$$

or equivalently, the following optimisation problem:

$$\min_x f(x) := \frac{1}{2}x^T Ax - b^T x$$

where, A is a symmetric and positive definite matrix.

Conjugate Gradient Method

Solution: Existence and Uniqueness

We know from elementary Linear Algebra that a linear system of equations $Ax = b$:

- Has unique solution if A is non-singular.
- Has infinite number of solutions if under-determined system.
- Has no solution if over-determined system (can have solution under certain conditions).

Conjugate Gradient Method by design assumes that A is symmetric and positive definite. i.e. the existence and uniqueness is guaranteed by design.

Conjugate Gradient Method

Error Estimate

UPPSALA UNIVERSITET

The error estimate for CG method is bound at iteration k as follows:

$$\left\| e^{(k)} \right\|_A \leq \frac{2C^k}{1 + C^{2k}} \left\| e^{(0)} \right\|_A \qquad \text{where } C = \frac{\sqrt{K_2(A)} - 1}{\sqrt{K_2(A)} + 1}$$

Here $K_2(A) = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}$ is the condition number of matrix A.

$\lambda_{\max}(A)$ and $\lambda_{\min}(A)$ are the largest and smallest eigenvalues of matrix A respectively.

Conjugate Gradient Method

Convergence Rate

From the error estimate of CG method at step k we have

$$\left\|e^{(k)}\right\|_A \leq \frac{2C^k}{1+C^{2k}} \left\|e^{(0)}\right\|_A$$

$$\Longrightarrow \frac{\left\|e^{(k+1)}\right\|_A}{\left\|e^{(k)}\right\|_A} \leq \frac{C^{k+1}}{1+C^{2(k+1)}} \frac{1+C^{2k}}{C^k} = C \underbrace{\frac{1+C^{2k}}{1+C^{2(k+1)}}}_{\text{goes to 1}}$$

$$\Longrightarrow \lim_{k\to\infty} \frac{\left\|e^{(k+1)}\right\|_A}{\left\|e^{(k)}\right\|_A} \leq C$$

Hence the rate of convergence of CG method is C.

Conjugate Gradient Method

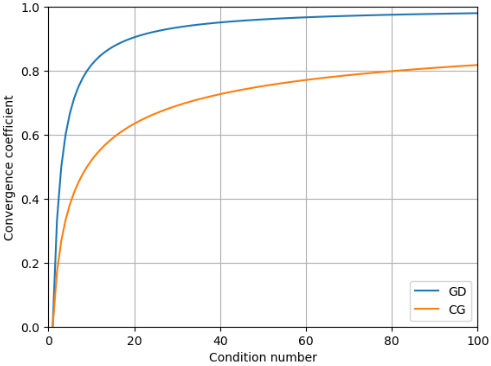
Convergence Rate - CG vs. GD

For CG method:

$$w \le \frac{\sqrt{K_2(A)} - 1}{\sqrt{K_2(A)} + 1}$$

For GD method:

$$w \le \frac{K_2(A) - 1}{K_2(A) + 1}$$



Conjugate Gradient Method

Convergence Rate - Discussion

The convergence rate derived in previous slide is the worst case analysis. In practice we have faster convergence because (after some messy derivations):

$$\left\|e^{(k)}\right\|_A^2 = \min_{P_k} \sum_{i=1}^{n} \lambda_i [1 + \lambda_i P_k(\lambda_i)]^2 \xi_i^2$$

Here λ_i's are the Eigenvalues of matrix A, and ξ_i are coefficients such that $x_0 - x^* = \sum_{i=1}^{n} \xi_i v_i$, where v_i are eigenvectors corresponding to eigenvalues of A.

Conjugate Gradient Method

Method Discussion

CG method works by generating a set of conjugate vectors. A set of vectors $\{p_0, p_1, \ldots, p_k\}$ is conjugate with respect to matrix A if:

$$p_i^T A p_j = 0 \qquad \forall i \neq j$$

At step k: CG method first finds a search direction p_k to optimize along by computing a conjugate vector. Then uses the residual information r_k at this step to compute the update along this particular direction.

Conjugate Gradient Method

Discussion Subspace and Norms

Each search direction p_k and residual r_k generated by CG method is contained in *Krylov subspace of degree k for* r_0, defined as:

$$\mathscr{K}(r_0; k) := \text{span}\{r_0, Ar_0, \dots, A^k r_0\}$$

This piece of information is crucial and needed to prove the convergence rate of Conjugate Gradient Method.

Conjugate Gradient Method

Convergence Proof

From the update rule of CG method we have vector x at step $k+1$ as follows:

$$x_{k+1} = x_0 + \alpha_0 p_0 + \alpha_1 p_1 + \cdots + \alpha_k p_k$$

If $x_{k+1} \neq x^*$ then the following property holds (why?):

$$\mathrm{span}\{p_0, p_1, \ldots, p_k\} = \mathrm{span}\{r_0, Ar_0, \ldots, A^k r_0\}$$

Consequently, we have:

$$\begin{aligned} x_{k+1} &= x_0 + \gamma_0 r_0 + \gamma_1 A r_0 + \cdots + \gamma_k A^k r_0 \\ &= x_0 + \underbrace{\left(\gamma_0 \mathbb{I} + \gamma_1 A + \cdots + \gamma_k A^k\right)}_{P_k^*(A)} r_0 = x_0 + P_k^*(A) r_0 \end{aligned}$$

Conjugate Gradient Method

Convergence Proof - continued

Before we proceed further we first state a theorem (and leave out the proof).

Theorem: Let $x_0 \in \mathbb{R}^n$ be the initial guess and $\{x_k\}$ then x_k is the minimizer of $f(x) = \frac{1}{2}x^TAx + b^Tx$ over the set

$$\{x | x = x_0 + \operatorname{span}\{p_0, p_1, \ldots, p_{k-1}\}\}$$

From the definition of A-norm, the function $f(x)$ and the fact that x^* minimizes $f(x)$ it can be shown that:

$$\|x - x^*\|_A^2 = \frac{1}{2}(x - x^*)^T A(x - x^*) = f(x) - f(x^*)$$

Conjugate Gradient Method

Convergence Proof - continued

Since x_{k+1} minimizes function $f(x)$, hence it also minimizes $\|x - x^*\|_A^2$ over the set $x_0 + \text{span}\{p_0, p_1, \ldots, p_{k-1}\}$ which is the same as $x_0 + \text{span}\{r_0, Ar_0, \ldots, A^k r_0\}$.

It follows from $x_{k+1} = x_0 + P_k^*(A)r_0$ that the polynomial P_k^* solves the following problem in which minimum is taken over the space of all possible polynomials of degree k:

$$\min_{P_k} \|x_0 + P_k(A)r_0 - x^*\|_A$$

Conjugate Gradient Method

Convergence Proof - continued

Using the optimality property (i.e. $Ax^* = b$) we have:

$$r_0 = Ax_0 - b = Ax_0 - Ax^* = A(x_0 - x^*)$$

Consequently,

$$x_{k+1} - x^* = x_0 + P_k(A)r_0 - x^* = [\mathbb{I} + P_k(A)A]\,(x_0 - x^*)$$

Eigendecomposition of symmetric matrix A gives us $A = \sum_{i=1}^{n} \lambda_i v_i v_i^T$.
Since the eigenvectors span the whole space $\mathbb{R}^n$, we can write:

$$x_0 - x^* = \sum_{i=1}^{n} \xi_i v_i$$

Conjugate Gradient Method

Convergence Proof - continued

Eigenvectors of matrix A are also eigenvectors of the polynomial of matrix A i.e. $P_k(A)v_i = P_k(\lambda_i)v_i, \ \forall i = 1, 2, \ldots, n$. As a result we get:

$$x_{k+1} - x^* = \sum_{i=1}^{n} [1 + \lambda_i P_k^*(\lambda_i)]\xi_i v_i$$

Since $\|z\|_A^2 = z^T A z = \sum_{i=1}^{n} \lambda_i (v_i^T z)^2$, we have

$$\|x_{k+1} - x^*\|_A^2 = \sum_{i=1}^{n} \lambda_i [1 + \lambda_i P_k^*(\lambda_i)]^2 \xi_i^2$$

References

[1] Canuto, C., Hussaini, M. Y., Quarteroni, A., and Zang, T. A.
Spectral Methods.
Springer Berlin Heidelberg, 2006.

[2] Nocedal, J., and Wright, S. J.
Numerical Optimization, *2nd ed*.
Springer, New York, NY, USA, 2006.

[3] Quarteroni, A., Sacco, R., and Saleri, F.
Numerical Mathematics.
Springer New York, 2007.

UPPSALA
UNIVERSITET

UPPSALA
UNIVERSITET

Thank you for your attention!