Quantitative Criteria and Redundancy in Bayesian Automatic Adaptive Quadrature

Gh. Adam[1,2], S. Adam[1,2]

e-mail: adamg@jinr.ru, [1]Laboratory of Information Technologies, JINR, Dubna
[2]DFT IFIN-HH, Magurele – Bucharest, 077125, Romania

1. Introduction

The Bayesian automatic adaptive quadrature (Baaq) [1, 2] of the Riemann integral

$$I \equiv I_{ab}[f] = \int_a^b f(x)dx, \quad -\infty < a < b < +\infty \quad (1)$$

yields a pair $\{Q, E > 0\}$ representing a globally adaptive numerical solution within the non-negative input accuracy specifications $\{\varepsilon_r > 0, \varepsilon_a \geq 0\}$, hopefully satisfying the termination criterion,

$$|I - Q| < E < \max\{\varepsilon_r|I|, \varepsilon_a\} \approx \max\{\varepsilon_r|Q|, \varepsilon_a\}.$$

To this aim, the quantities Q and E are computed merging rigorous mathematical criteria for the control of the output quality with the reality of the hardware and software environments, within an approach based on two pillars:

(1) A *subrange subdivision strategy* of the integration domain $[a, b]$ implementing the ordering of the generated subranges into a *priority queue*.

(2) Sets of convenient *local quadrature rules* which yield, over each subrange $[\alpha, \beta] \subseteq [a, b]$, a *local pair* $\{q, e\}$ for the *quadrature sum* q and its associated *local error estimate* $e > 0$, respectively.

The global pair terms Q and E are obtained as sums of q and e over subranges.

The present report summarizes the progress achieved along two distinct lines:

(i) The derivation of quantitative conditioning criteria establishing admissible rates of variation of the integrand $f(x)$ inside monotonicity intervals.

(ii) The possibility to exploit the opportunities offered by the multi-core computing architectures for sharpening the local error estimates over subranges.

2. Quantitative conditioning criteria over strict monotonicity intervals

One of the reasons preventing the derivation of a numerical local quadrature rule output within prescribed accuracy stems from the occurrence, over the strict monotonicity intervals, of fast integrand variations which are in excess of the bounds following from those enforced by the interpolatory polynomials spanning the involved quadrature sums.

There are two pieces of information which enter the derivation of conditioning criteria enforcing the fulfillment of such bounds by subrange subdivision:

- The *integrand profile* over the integration (sub)range of interest $[\alpha, \beta] \subseteq [a, b]$, which collects the computed integrand values over this range,

$$\{(x_\mu, f_\mu), \mu = 0, \cdots, M | \alpha = x_0 < \cdots < x_M = \beta\} \quad (2)$$

- The *floating point degree of precision* (fpdp), d_{fp}, which supersedes the usually assumed *algebraic degree of precision*, d, of a quadrature sum [3, 4].

Given the discrete integrand sampling (2), *strict monotonicity intervals* $[x_\lambda, x_{\lambda+\Lambda}]$,

$$x_\lambda < x_{\lambda+1} < \cdots < x_{\lambda+\Lambda} \quad (3)$$

are defined from the condition of the *sign constancy* of the floating point differences $\delta_{k,k+1} = f_{k+1} - f_k$ which give the integrand variation inbetween successive abscissas x_k and x_{k+1}, i.e.,

$$\delta_{k-1,k} \cdot \delta_{k,k+1} > 0, \ \forall k \in \{\lambda+1, \cdots, \lambda+\Lambda-1\}.$$

The left end $x_\lambda > x_0 = a$ of the integrand profile (2) defines a *local extremum* of $f(x)$ provided $\delta_{\lambda-1,\lambda} \cdot \delta_{\lambda,\lambda+1} < 0$. Then we say that $[x_\lambda, x_{\lambda+\Lambda}]$ is *closed to the left*.

Otherwise, if $\delta_{\lambda-1,\lambda} \cdot \delta_{\lambda,\lambda+1} = 0$, then there is a *plateau* of $f(x)$ to the left of x_λ and we say that $[x_\lambda, x_{\lambda+\Lambda}]$ is *open to the left*.

If $x_\lambda = x_0 = a$, then the integrand function $f(x)$ is undefined at $x < a$. In this case, we also say that the strict monotonicity interval $[x_0, x_\Lambda]$ is open to the left.

Similar definition hold with respect to the right end $x_{\lambda+\Lambda}$.

With this preparative at hand, we characterize the sequence (3) as follows:
– *closed* if both x_λ and $x_{\lambda+\Lambda}$ are extrema of $f(x)$;
– *open* if neither x_λ nor $x_{\lambda+\Lambda}$ are extrema of $f(x)$;
– *closed to the left* if x_λ is an extremum, while $x_{\lambda+\Lambda}$ is not;
– *closed to the right* if $x_{\lambda+\Lambda}$ is an extremum, while x_λ is not.

There are three kinds of quantities which enter, directly or indirectly, the conditioning criteria:
1. The *first order divided difference*,

$$d_{k-1,k} = \delta_{k-1,k}/(x_k - x_{k-1}), \quad (4)$$

which approximates the *slope* of $f(x)$ over (x_{k-1}, x_k).

2. The *second order divided difference*,

$$d_{k-1,k,k+1} = (d_{k,k+1} - d_{k-1,k})/(x_{k+1} - x_{k-1}), \quad (5)$$

which approximates the *curvature* of $f(x)$ over (x_{k-1}, x_{k+1}).

3. The *max-norm bound*, ν_{B}, which defines a characteristic *sensitivity threshold* of the Bayesian inferences over strict monotonicity intervals,

$$
\begin{aligned}
\nu_{\mathrm{B}} &= \max\{\nu_{\mathrm{d}}, \nu_{\mathrm{fp}}\}, & (6)\\
\nu_{\mathrm{d}} &= \max_{1 \le \mu \le M}(x_\mu - x_{\mu-1})\\
\nu_{\mathrm{fp}} &= (\beta - \alpha)/(d_{\mathrm{fp}} - 1).
\end{aligned}
$$

Bayesian inferences on the integrand conditioning over (3) can be drawn provided its ends are *isolated* from each other. This condition is assumed to be satisfied provided it is possible to define over (3) two second order divided differences $d_{\gamma-2,\gamma-1,\gamma}$ and $d_{\delta,\delta+1,\delta+2}$ such that $\delta \ge \gamma$. This definition results in the condition $\Lambda \ge \lambda + 4$, i.e., that the strict monotonicity interval extends over five consecutive abscissas at least.

The derivation of quantitative conditioning criteria over the strict monotonicity interval (3) is enabled by the following *well-conditioning Ansatz*:

<u>Ansatz.</u> Let $\{x_{k-1} < x_k < x_{k+1}\}$ denote an ordered triplet inside the set with isolated endpoints (3).

(a) If $x_{k+1} - x_{k-1} < \nu_{\mathrm{B}}$, then the expected rate of variation of a well-conditioned integrand over $[x_{k-1}, x_{k+1}]$ cannot exceed that of a second degree polynomial.

(b) If $x_{k+1} - x_{k-1} \ge \nu_{\mathrm{B}}$, then this rate of variation cannot exceed that of a third degree polynomial.

The use of this well-conditioning Ansatz enabled the derivation [5] of several consistency criteria involving distinctly the inner points and the endpoints of the strict monotonicity interval (3).

For Baaqs spanned by orthogonal polynomials, the statement (a) of the Ansatz covers the densely sampled regions around the ends of every subrange $[\alpha, \beta] \subseteq [a, b]$, while the statement (b) covers the sparsely sampled regions around the centre of $[\alpha, \beta] \subseteq [a, b]$.

A straightforward consequence of this well-conditioning Ansatz is the enforcement of two-sided bounds to the ratios of the slopes of the integrand over (x_{k-1}, x_k) and (x_k, x_{k+1}) inside each triplet of interest. Let

$$D_k = |d_{k,k+1}|/|d_{k-1,k}|, \ \rho_k = (x_{k+1} - x_k)/(x_k - x_{k-1}).$$

Inside a *densely sampled region*, the integrand is well-conditioned provided one of the following two criteria is satisfied:

(a1) If $D_k > 1$ then $1 < D_k < 2 + \rho_k$.

(a2) If $D_k < 1$ then $1 < D_k^{-1} < 2 + \rho_k^{-1}$.

Inside a *sparsely sampled region*, the integrand is well-conditioned provided one of the following two criteria is satisfied:

(b1) If $D_k > 1$ then $1 < D_k < 3 + 3\rho_k + \rho_k^2$.

(b2) If $D_k < 1$ then $1 < D_k^{-1} < 3 + 3\rho_k^{-1} + \rho_k^{-2}$.

3. Sharpening local error estimates using redundancy in the Baaq

For almost six decades, the main stream of the hardware development for scientific computing essentially followed the von Neumann model based on a single central processing unit (CPU). The sequential algorithms developed for the von Neumann computer were implemented into software reaching a high level of optimization.

The last decade marked, however, an essential shift of paradigm, such that the newly produced hardware for scientific computing consists nowadays exclusively of multi-core processors, possibly supplemented with GPU accelerators.

The Baaq approach to the numerical solution of the integrals can get substantial benefit from the possibilities offered by the new hardware architecture to improve both the algorithm accuracy and the speed of computations. The recent proposal [6] dealt with a modification of the paradigm concerning the implementation of the local quadrature rules using redundancy.

Within the previous approach to the automatic adaptive quadrature [7], in order to get the pair $\{q, e\}$ over the current subrange $[\alpha, \beta] \subseteq [a, b]$, two embedded symmetric quadrature sums, $q_{2n,\alpha\beta}[f]$ and $q_{n,\alpha\beta}[f]$, are used. Then the more accurate approximation $q_{2n,\alpha\beta}[f]$ is taken for the value of the integral $I_{\alpha\beta}[f]$, while the modulus of the reminder $|I_{\alpha\beta}[f] - q_{2n,\alpha\beta}[f]|$ is estimated from the difference $|q_{2n,\alpha\beta}[f] - q_{n,\alpha\beta}[f]|$ using *probabilistic/heuristic* arguments. On one side, there is compelling empirical evidence that, for most problems asking the solution of (1), the obtained local error estimates are too conservative. On the other side, the fact that the quadrature knots of $q_{n,\alpha\beta}$ are a subset of $q_{2n,\alpha\beta}$ is an important source of pitfalls concerning the reliability of the inferences based on the use of such an error estimate, which was amply documented in the literature.

The new proposal for the Baaq local error estimate starts with two unrelated to each other quadrature sums over $[\alpha, \beta] \subseteq [a, b]$:

-
$$q_{\mathrm{CC},\alpha\beta}[f] = \int_\alpha^\beta p_{\mathrm{CC},\alpha\beta}(x)\, dx$$

where $p_{\mathrm{CC},\alpha\beta}(x)$ is the interpolatory polynomial which equates the integrand $f(x)$ at

the reduced Clenshaw-Curtis (CC) quadrature knots defined as the extrema of the Chebyshev polynomial of the first kind $T_{32}(\xi)$ over $[-1, 1]$;

•

$$q_{\mathrm{GK},\alpha\beta}[f] = \int_\alpha^\beta p_{\mathrm{GK},\alpha\beta}(x) \, dx$$

where $p_{\mathrm{GK},\alpha\beta}(x)$ is the interpolatory polynomial which equates $f(x)$ at the reduced Gauss-Kronrod (GK) 10-21 quadrature knots over $(-1, 1)$ [7].

The local quadrature rule pair $\{q, e > 0\}$ is then obtained from q_{CC} and q_{GK} guided by the following considerations.

The choice

$$q = q_{\mathrm{CC},\alpha\beta}[f] \tag{7}$$

is made based on the observation that the interpolatory polynomial $p_{\mathrm{CC},\alpha\beta}(x)$ lies close to the minimax approximating polynomial of $f(x)$ everywhere inside $[\alpha, \beta]$, while the deviation of the interpolatory polynomial $p_{\mathrm{GK},\alpha\beta}(x)$ from $f(x)$ may get large inbetween the successive quadrature knots. Moreover, the algebraic degree of precision of q_{CC}, $d_{\mathrm{CC}} = 32$, is slightly larger than that of q_{GK}, $d_{\mathrm{GK}} = 31$.

The choice

$$e = \max\{|q_{\mathrm{CC},\alpha\beta}[f] - q_{\mathrm{GK},\alpha\beta}[f]|, \varepsilon_0 \cdot q_{\mathrm{tr},\alpha\beta}[|f|]\}, \tag{8}$$

is made based on the *statistical independence* of the quadrature knots of q_{CC} and q_{GK}, which share a single common abscissa, the centre $\gamma = (\beta + \alpha)/2$ of $[\alpha, \beta]$. The last term entering (8) is the product between ε_0, the machine epsilon with respect to addition, and $q_{\mathrm{tr},\alpha\beta}[|f|]$, the integral over $[\alpha, \beta]$ of $|f(x)|$, the modulus of $f(x)$, computed by means of the trapeze rule using the mesh consisting of the union of the quadrature knots of q_{CC} and q_{GK} over $[-1, 1]$.

Within a multi-core hardware environment, the computation of $f(x)$ at the CC and GK abscissas, as well as of the quadrature sums q_{CC}, q_{GK}, q_{tr} may be done independently on different cores, such that the redundant addition of supplementary quadrature knots does not result into sizable increase of the computing time of the local quadrature rule pairs $\{q, e\}$ over subranges.

Two main advantages are immediate. First, the expression (8) secures a *substantial sharpening of the local error estimates*, hence quicker end of the global computation. Second, this secures a *substantial enrichment* of the statistics enabling Bayesian inferences on the integrand conditioning over the subranges characterized by slow convergence of the local quadrature rules under subrange subdivisions.

References

[1] Gh. Adam, S. Adam, *Principles of the Bayesian automatic adaptive quadrature*, Numerical Methods and Programming: Advanced Computing, **10**, (2009) 391–397 (http://num-meth.srcc.msu.ru).

[2] Gh. Adam, S. Adam, *Bayesian automatic adaptive quadrature: an overview*, in Mathematical Modeling and Computational Science, MMCP 2011, LNCS, vol. 7125, Gh. Adam, J. Buša, M. Hnatič, Eds. Heidelberg: Springer, 2012, pp. 1-16.

[3] S. Adam, Gh. Adam, *Floating point degree of precision in numerical quadrature*, in Mathematical Modeling and Computational Science, MMCP 2011, LNCS, vol. 7125, Gh. Adam, J. Buša, M. Hnatič, Eds. Heidelberg: Springer, 2012, pp. 189-194.

[4] S. Adam, Gh. Adam, *The Concept of Floating Point Degree of Precision*, LIT-JINR Scientific Report 2010–2011.

[5] Gh. Adam, S. Adam, *Quantitative Conditioning Criteria in Bayesian Automatic Adaptive Quadrature*, in Grid, Cloud & High Performance Computing Science (RO-LCG2012), U.T. Press, Cluj Napoca, 2012, pp. 35–38.

[6] Gh. Adam, S. Adam, *Sharpening local error estimates using redundancy in Bayesian automatic adaptive quadrature*, Oral Communication, MMCP 2013 International Conference, JINR-Dubna, July 8–12, 2013, Book of Abstracts, pp. 23–24, http://mmcp2013.jinr.ru/files/boa.pdf

[7] R. Piessens, E. deDoncker-Kapenga, C.W. Überhuber, D.K. Kahaner, *QUADPACK, a subroutine package for automatic integration*, Springer Verlag, Berlin (1983).