The Bayesian Approach to Numerical Problems: Results for Zero Finding

Erich Novak

Mathematisches Institut, Universität Erlangen-Nürnberg
Bismarckstr. 1 1/2, 91054 Erlangen, Germany, novak@mi.uni-erlangen.de

Abstract. We explain the Bayesian approach to nonlinear numerical problems. In particular, we present average case results for zero finding. We construct a hybrid enclosing method that is almost optimal in the average case sense. Our results differ significantly from known results in the asymptotic setting or in the worst case setting. This paper can serve as an introduction to a recent more theoretical paper of Novak, Ritter, and Woźniakowski. We also present numerical test results and comparisons with other zero finding methods.

1. Results for zero finding

We consider classes F_r of functions defined by

$$F_r = \{f \in C^r[0,1] \mid f(0) < 0 < f(1)\},$$

where $r \in \mathbf{N}_0$. We study zero finding methods $\widetilde{S}$ which are based on function values $f(x_i)$ or derivatives $f^{(k_i)}(x_i)$ at adaptively (sequentially) chosen knots x_i. Of course we assume that $k_i \le r$. The output is an interval $[s,t]$ such that $f(s) \cdot f(t) \le 0$. The error is defined by the criterion

$$\Delta(f, \widetilde{S}) = t - s, \tag{1}$$

the worst case error of $\widetilde{S}$ is defined by

$$\Delta_{\max}(\widetilde{S}) = \sup_{f \in F_r} \Delta(f, \widetilde{S}).$$

We present a special algorithm S_ε with $\Delta_{\max}(S_\varepsilon) \le \varepsilon$, also for $r = 0$. By the error criterion (1) it is *guaranteed* that we find a zero with error bounded by $\Delta_{\max}(\widetilde{S})$, hence only *enclosing methods* can be used. Neglecting the arithmetic cost which is caused, for instance, by the selection of the knots we define

$$\mathrm{cost}(f, \widetilde{S}) = n(f),$$

where $n(f)$ is the number of function or derivative evaluations used to compute $\widetilde{S}(f)$. Error and cost are important quantities in a comparison of numerical methods. Such a comparison can be made in the asymptotic setting, in the worst case setting, and in the average case setting. See also the survey paper [12].

In the *asymptotic setting* we ask, for example, what order of convergence can be achieved. The error estimates contain a constant c_f that depends on f and usually is unknown. Therefore we do not have a very good control over the cost that is needed to compute a zero up to a given error ε. We also note that in the asymptotic setting often only smooth functions with simple zeros are studied.

In the *worst case setting* we compare the maximal error $\Delta_{\max}(\widetilde{S})$ with the maximal cost

$$\mathrm{cost}_{\max}(\widetilde{S}) = \sup_{f \in F_r} \mathrm{cost}(f, \widetilde{S}).$$

It is known that the bisection method is optimal for all classes F_r and also for many other classes, such as

$$F^* = \{f \in C^\infty[0,1] \mid f(0) < 0 < f(1),\ f' > 0\}.$$

This may be surprising since iterative methods are known which are globally and superlinearily convergent for smooth functions having simple zeros. However, asymptotic results do not imply error bounds which hold after a finite number of steps.

The results explained in this paper show that an additional *average case analysis* is of interest in particular for nonlinear problems like zero finding. Here we consider the average cost

$$\mathrm{cost}_{\mathrm{aver}}(\widetilde{S}) = \int \mathrm{cost}(f, \widetilde{S})\, dP(f)$$

with respect to a probability measure P on F_r. The latter quantity takes into account also those functions f where $\widetilde{S}$ performs well, and average case results should match practical experience more closely for reasonably chosen measures P.

First we describe our method S_ε. We assume that an error bound $\varepsilon > 0$ is given and that we can evaluate functions $f \in F_0$.

$i := 0$;
$[s_0, t_0] := [0,1]$; compute $f(0),\ f(1)$;
do
  for $j := 1, 2$ do (regula falsi)
    $i := i + 1$;
    $x_i :=$ SEC(s_{i-1}, t_{i-1}); compute $f(x_i), [s_i, t_i]$;
  od;
  repeat (secant)
    $q :=$ SEC(x_i, x_{i-1});
    if $q \notin\ ]s_i, t_i[$ then break fi;
    $i := i + 1$;
    $x_i := q$; compute $f(x_i)$, $[s_i, t_i]$;
  until $t_i - s_i > (t_{i-3} - s_{i-3})/2$;
  $i := i + 1$; (bisection)
  $x_i := (s_i + t_i)/2$; compute $f(x_i)$, $[s_i, t_i]$;
od;

After each computation of a function value $f(x_i)$ we compute the new enclosing interval $[s_i, t_i]$. We stop if $t_i - s_i \le \varepsilon$. In that case we return $s = s_i$ and $t = t_i$.

Remarks. We mention some properties of the method S_ε that are also important for the proof of average case results. The method uses steps of the regula falsi (R), the secant method (S), and the bisection method (B). A typical pattern is R R S ... S B R R S B R R S S S S S S ... We always have a length reduction $t_i - s_i \le (t_{i-4} - s_{i-4})/2$ of the interval with a guaranteed zero. The error of the method is bounded by ε for each $f \in F_0$. The computational cost is proportional to the number of function evaluations. Enclosing methods were recently studied in [1] and [14].

We deal with infinite dimensional function classes F, and therefore no analogue to the Lebesgue measure is available. So far Gaussian measures, which enjoy certain invariance properties, have been used in most of the papers dealing with the average case on such classes F.
On classes like F_r we use conditional r-fold Wiener measures P, see [12] and [13]. These Gaussian measures are derived from the classical Wiener measure by r-fold integration and translation by suitable polynomials. The following result holds for r-fold Wiener measures and $r \ge 2$. The constant c in (2) depends on the measure P, the proof in [13] gives a suitable c that can be computed numerically.

Theorem 1. Let $0 < \varepsilon < 1/2$. For the hybrid secant-bisection method S_ε,

$$\mathrm{cost}_{\mathrm{aver}}(S_\varepsilon) \le \frac{1}{\log \beta} \cdot \log\log(1/\varepsilon) + c, \tag{2}$$

where $\beta = \frac{1+\sqrt{5}}{2}$.

We only sketch the basic idea of the rather technical and long proof. For $\gamma_1, \gamma_2, \gamma_3 > 0$ we define a subset G of F_r by

$$G = \{g \in F_r \mid |g''(x) - g''(y)| \le \gamma_1\, |x - y|^{1/3},\ |g''(z)| \ge \gamma_2,\ |g'(z)| \ge \gamma_3 \text{ if } g(z) = 0\}.$$

Let $g \in G$. After $k = k(G)$ steps we can guarantee the following: If there is a further bisection step then this is the last one and we obtain the pattern B R R S S S S ... It is possible to give explicit estimates of $k(G)$ by a worst case analysis on the class G. It is also possible to estimate the probability of G and so, finally, the result can be proved.
One should note that such an average case analysis includes a worst case analysis for many different subsets G of F_r. It is crucial, of course, that a membership $f \in G$ can not be checked by any algorithm that is based on a finite information. Hence the constants γ_i cannot be used in a program.

We claim that the algorithm S_ε is almost optimal in a very strong sense. Actually we can prove a lower bound for very general methods: instead of function evaluations at any knots we also allow the evaluation of derivatives; instead of enclosing methods with a guaranteed error ε we consider methods with output $x \in [0, 1]$ and define now the error by the distance to the next zero of f. In addition, we only require that the average error of a method is bounded by ε. Nevertheless, the average cost of any such algorithm cannot be much smaller compared to the cost of our algorithm S_ε. Again we consider r-fold Wiener measures, where now $r \in \mathbf{N}_0$ is arbitrary.

Theorem 2. Let $\varepsilon > 0$ and assume that $\widetilde{S}$ is a method with

$$\int \inf\{|\widetilde{S}(f) - x| \mid f(x) = 0\} \, dP(f) \le \varepsilon.$$

Then

$$\mathrm{cost}_{\mathrm{aver}}(\widetilde{S}) \ge \frac{1}{\log \alpha} \cdot \log\log(1/\varepsilon) + c_\alpha,$$

for any α with $\alpha > r + 1/2$.

The constant c_α depends on the measure, but is of course independent of $\widetilde{S}$. We believe that the bound $r+1/2$ is optimal. Together with Theorem 1 we obtain that our method S_ε is almost optimal if $r \ge 2$. Possibly, this particular method is also optimal for $r = 0$ and $r = 1$ but we do not know good bounds in these cases.
The stopping rule "$t_i - s_i \le \varepsilon$" is *adaptive* since the number $n(f)$ of function evaluations depends on $f \in F_r$. Such an adaptive stopping rule is crucial. It is known that with a nonadaptive stopping rule

$$n(f) = n \quad \text{for all } f \in F_r$$

one can only achieve a linear convergence, i.e., there is a lower bound

$$\Delta_{\mathrm{aver}}(\widetilde{S}) \ge \alpha^n,$$

see [10] and [17]. Note the huge difference between the orders $\log(1/\epsilon)$ and $\log\log(1/\epsilon)$ of the worst and average case complexity. This difference is due to the fact that we switch from maximal cost to average cost, the difference between worst case errors and average case errors turns out to be a minor one for this particular problem. In Section 3 we present numerical examples for S_ε and compare the new method with other methods.

2. The Bayesian approach to numerical problems

We compare different numerical methods by their error and their cost. Usually such a comparison is made on a class of functions which might be defined, for instance, by smoothness properties. The overall performance of a method is very often understood in a worst case sense and the maximal error and cost are studied. An additional average case analysis is of interest in particular for nonlinear problems like global optimization or zero finding. It turns out that for nonlinear problems, as a rule, optimal average case methods differ significantly from optimal worst case methods.
For instance, the error of a method $\widetilde{S}$ for global optimization may be defined by

$$\Delta(f, \widetilde{S}) = \sup_x f(x) - f(\widetilde{S}(f)).$$

If $\widetilde{S}$ uses only function values to determine a point $\widetilde{S}(f)$ where f is large, then the total number $n(f,\widetilde{S})$ of evaluations serves as a crude definition of the cost. The complexity

$$\mathrm{comp}(\epsilon) = \inf\{\mathrm{cost}(\widetilde{S}) : \Delta(\widetilde{S}) \le \epsilon\}$$

of a numerical problem is the minimal cost which is needed to obtain an error ϵ. Here cost and error are defined in the worst case or average case sense. Strictly speaking, the definition of cost requires a model of computation. This model determines the kind of information about f which is computable and the operations to combine the information, together with the respective cost, see [9] and [20].

For linear problems, like integration or reconstruction of functions, and Gaussian measures general results say that adaption and adaptive stopping rules do not help and linear algorithms are optimal, see [20] and [21]. In several papers r-fold Wiener measures are used to study linear problems like integration or reconstruction of functions, see, e.g., [22].

Some nonlinear problems were studied under probabilistic assumptions by several authors and upper bounds are obtained. A partial list includes [16] and [19] on zero finding for (systems of) polynomials, [3] on linear programming, and [5] on integral equations. The problem of numerical integration of monotone functions was studied in [8]. For all these problems, however, tight lower bounds are still unknown.

Consider the problem of global optimization on any convex and symmetric class F of functions. We study methods $\widetilde{S}$ which evaluate $f \in F$ at adaptively chosen knots. The worst case complexity of global optimization and reconstruction in L_∞-norm are closely related. Global optimization is roughly as hard as reconstruction on the whole domain, although subregions may sometimes be excluded during a computation, for instance, in the case of Lipschitz functions. Furthermore, adaption does not help for global optimization in the worst case setting, see [7].

Consider, however, the class $F = C[0,1]$ equipped with the Wiener measure P. The error of any nonadaptive method which uses n function values is of order at least $n^{-1/2}$ and this order is achieved with equidistant knots, see [17]. Recently Calvin [4] constructed adaptive methods that are much better. Thus adaption turns out to be very powerful in the average case setting. Furthermore, on the average, optimization is significantly easier than reconstruction.

Constructing and analyzing good algorithms for global optimization of multivariate functions is a challenging task. In an average case analysis one could use the Wiener sheet measure on $F = C[0,1]^d$. For the linear problem of reconstruction in L_2-norm the complexity is known and good algorithms are based on hyperbolic cross points, see [22]. An adaptive optimization method based on these points is constructed in [11], and numerical results are given that are quite promissing. This method might be an interesting candidate for an average case analysis.

3. Numerical examples for zero finding methods

The main purpose of our study [13] was to understand the different behavior of optimal methods in the different settings: asymptotic, worst case, and average case.

The method S_ε that we presented in Section 1 allows a theoretical analysis. It can be expected that more sophisticated methods yield better results for sufficiently nice functions. For smooth functions with a simple zero it is better to use a higher order interpolation instead of linear interpolation. In spite of these reservations, we made some numerical tests and compared S_ε with some well known methods. Our new method performed pretty well and the theoretical insight is confirmed by the numerical results.

To get a clear picture of the pros and cons of the different methods, it is necessary to use many different test functions with different properties. As it can be expected, there is no method that works equally well for all possible functions. For our comparison we used 31 test functions. These functions were collected by Thomas Lindner, some of them were used earlier in the literature. We thank Thomas Lindner as well as Martin Ruprecht and Matthias Kutschenreuter who did parts of the tests. We always use functions $f \in C([0,1])$ with $f(0) \cdot f(1) < 0$. The first two functions are given by

$$f_1(x) = 5x - e^x \quad \text{and} \quad f_2(x) = \sin(x - 1/3).$$

The next three functions are

$$f_{3,n}(x) = e^{nx} - e^{2n/7}$$

with $n = 2$, 5, and 10. Three functions are given by

$$f_{4,n} = 2xe^{-n} + 1 - 2e^{-nx},$$

where $n = 2$, 10, and 100. Three functions are

$$f_{5,n}(x) = (nx - n/5 - 1/2) \cdot (x + n - 1/5)/(n-1),$$

where $n = 2$, 12, and 1000. Three functions are given by

$$f_{6,n}(x) = (1 + (1-n)^2)x - (1 - nx)^2,$$

with $n = 2$, 10, and 100. The next family is given by

$$f_{7,n}(x) = \frac{1 - 3x}{1 + nx},$$

we take $n = 2$, 10, 100, 1000, and 10000. We also take 5 functions of the family

$$f_{8,n}(x) = (x - 4/7)^3 - 1/(10n),$$

with $n = 2$, 10, 100, 1000, and 10000. All these functions are smooth and have a simple zero. Therefore the classical results in the asymptotic setting can be applied. We also used a family $f_{9,n}$ of functions with a multiple zero and a family $f_{10,n}$ containing functions that are not in C^1. The family $f_{9,n}$ is given by

$$f_{9,n}(x) = (x - 3/7)^{2n-1},$$

we take $n = 2$, 3, and 4. The last family finally is given by

$$f_{10,n}(x) = \begin{cases} (x - 7/9)(x + 1) & \text{if } x \le 7/9 \\ (x - 7/9)(x + n) & \text{else,} \end{cases}$$

we take $n = 2$, 10, 100, and 1000.

We tried the following methods: BIS (bisection), SEC (secant method), REG (regula falsi), NEW (Newton), HYP (simple hybrid method: one step of regula falsi alternating with one step of bisection), ILL (Illinois method, see [2]), PEG (Pegasus method, see [2]), ANB (Anderson, Björck, see [2]), PEK (Pegasus, King, see [6]), ANK (Anderson-Björck with King modification, see [6]), APS (Alefeld, Potra, Shi, see [1]), WDB (Wijngaarden, Dekker, Brent, see [15]), and NRW, the method S_ε that is described in Section 1.
We always started with $x_0 = 0$ and $x_1 = 1$. In case of the Newton method, where we only need one starting point, we used the endpoint with the smaller (in absolute value) function value.

We used two different stopping rules. In the first test series we stop if a function value with $|f(x)| \le \varepsilon = 10^{-13}$ was found. Some of the methods (SEC, REG, WDB) did not always stop in less than 500 steps, which is considered bad since the bisection method always stopped after less than 57 steps. There were only 5 methods that used less than 50 function values in all 31 cases: ANK, ANB, NRW, PEK, PEG.
We also used the root error criterion where we stopped if an enclosing interval of the length $\varepsilon = 10^{-13}$ was found. Some of the methods (SEC, NEW, REG) do not produce small enclosing intervals and therefore could not be used for the second stopping rule. The results were quite different for the families 1–8 of smooth functions with a simple zero and for the two other families. For families 1–8 the winners are WDB, ANB, APS, ANK, PEK, PEG, ILL using less than 25 function values for each function. If we also consider the families 9 and 10 of poorly behaved functions, then the winners are BIS, HYP, WDB, NRW, APS using less than 160 function values for each function.

No method was among the winners in each of the three categories. Often ANB and ANK are excellent, WDB was very good for the root error criterion. However, ANB and ANK had problems with multiple zeros (in case of the root error) and WDB was not always good with the residual error. We have the overall impression that our new method NRW belongs to the group of the better methods; it was the only method with a satisfactory behavior over all functions with respect to both error criteria.

References

[1] Alefeld, G., Potra, F. A., and Yixun Shi: On enclosing simple roots of nonlinear equations. *Math. of Comp.* 61 (1993), 733–744.

[2] Anderson, N., and Björck, Å.: A new high order method of regula falsi type for computing a root of an equation. *BIT* 13 (1973), 253–264.

[3] Borgwardt, K. H.: *The Simplex Method.* Springer-Verlag, Berlin, 1987.

[4] Calvin, J. M.: Average performance of a class of adaptive algorithms for global optimization. Preprint, Georgia Institute of Technology, Atlanta 1995.

[5] Heinrich, S.: Probalistic analysis of numerical methods for integral equations. *J. Integral Eq. Appl.* 3 (1991), 289–319.

[6] King R. F.: An improved Pegasus method for root finding. *BIT* 13 (1973), 423–427.

[7] Novak, E.: *Deterministic and Stochastic Error Bounds in Numerical Analysis.* Lect. Notes in Math. 1349, Springer-Verlag, Berlin 1988.

[8] Novak, E.: Quadrature formulas for monotone functions. *Proc. Amer. Math. Soc.* 115 (1992), 59–68.

[9] Novak, E.: The real number model in numerical analysis. *J. Complexity* 11 (1995), 57–73.

[10] Novak, E., and Ritter, K.: Average errors for zero finding: lower bounds. *Math. Z.* 211 (1992), 671–686.

[11] Novak, E., and Ritter, K.: Global optimization using hyperbolic cross points. To appear in: *Global Optimization*, C. A. Floudas and P. M. Pardalos (eds.), Kluwer, Dordrecht 1995.

[12] Novak, E., and Ritter, K.: Some complexity results for zero finding for univariate functions. *J. Complexity* 9 (1993), 15–40.

[13] Novak, E., Ritter, K., and Woźniakowski, H.: Average case optimality of a hybrid secant-bisection method. *Math. Comp.* 64 (1995), 1517–1539.

[14] Petković, M. S., and Herzberger, J.: Hybrid inclusion algorithms for polynomial multiple complex zeros in rectangular arithmetic. *Appl. Numer. Math.* 7 (1991), 241–262.

[15] W. H. Press et al.: *Numerical recipes in C.* Cambridge University Press, Cambridge 1988.

[16] Renegar, J.: On the efficiency of Newton's method in approximating all zeros of a system of complex polynomials. *Math. Oper. Res.* 12 (1987), 121–148.

[17] Ritter, K.: Approximation and optimization on the Wiener space. *J. Complexity* 6 (1990), 337–364.

[18] Ritter, K.: Average errors for zero finding: lower bounds for smooth or monotone functions. *Aequationes Math.* 48 (1994), 194–219.

[19] Shub, M., and Smale, S.: Complexity of Bezout's theorem. V: Polynomial time. *Theor. Comput. Sci.* 133 (1994), 141-164.

[20] Traub, J. F., Wasilkowski, G. W., and Woźniakowski, H.: *Information-Based Complexity*, Academic Press, New York 1988.

[21] Wasilkowski, G. W.: Information of varying cardinality. *J. Complexity* 2 (1986), 204–228.

[22] Woźniakowski, H.: Average case complexity of linear multivariate problems, Part 1: Theory, Part 2: Applications. *J. Complexity* 8 (1992), 337–372, 373–392.