Adaptive Step Size Random Search

MICHAEL A. SCHUMER, MEMBER, IEEE, AND KENNETH STEIGLITZ, MEMBER, IEEE

Abstract—**Fixed step size random search for minimization of functions of several parameters is described and compared with the fixed step size gradient method for a particular surface. A theoretical technique, using the optimum step size at each step, is analyzed. A practical adaptive step size random search algorithm is then proposed, and experimental experience is reported that shows the superiority of random search over other methods for sufficiently high dimension.**

INTRODUCTION

THE PROBLEM of locating the minimum of a function of several variables is one that arises frequently in many areas of technology, particularly in the design of adaptive control and communication systems.

The problem is: given the quality function $Q(\boldsymbol{X})$, where $\boldsymbol{X}$ is a vector of adjustable parameters $x_1, \cdots, x_n$; find the value of $\boldsymbol{X}$ that minimizes Q. The following assumptions are made.

1) Q is unimodal. If it is not, a global search can be carried out first to partition the parameter space into regions where Q is unimodal.
2) The structure of the function $Q(\boldsymbol{X})$ is completely unknown. The only way that information can be obtained is by evaluating Q at specific points. This means, for example, that derivatives of Q are not directly measurable (if they exist at all).
3) The only significant cost involved in the operation of a search procedure results from evaluating Q. Therefore, the fewer function evaluations required, the more desirable is the procedure.

Of course, any strategy suggested for solving the preceding problem can be evaluated only for a specific surface or class of surfaces. Analyses in this paper will be restricted to hyperspherical surfaces, and experimental results will be given for other surfaces as well.

Besides the many deterministic minimization algorithms developed, dating back to such classical methods as steepest descent, the Newton–Raphson method, and other gradient procedures, Brooks[1] and Rastrigin[2],[3] have suggested randomized search strategies.

Manuscript received October 30, 1967; revised January 19, 1968. This work was partially supported by Army Research Office—Durham under Contract DA-31-124-ARO-D-292, by a National Science Foundation Graduate Fellowship, and by a Research Grant from the Bendix Corporation. It made use of computer facilities supported in part by the National Science Foundation under Grant NSF-GP-579.

M. A. Schumer is with the Research Div., Raytheon Co., Waltham, Mass. 02154

K. Steiglitz is with the Dept. of Electrical Engineering, Princeton University, Princeton, N. J. 08540

Rastrigin has compared a fixed step size random search (FSSRS) method with a fixed step size gradient method and concluded that under certain circumstances FSSRS is superior. It is clear, however, that if the step size of the random search method were optimum at each step, even better performance would result. In this paper, a hypothetical random search method that uses the optimum step size at each point will be analyzed for a hyperspherical surface. An adaptive step size random search (ASSRS) method will then be proposed that approximates the performance of the optimum step size random search (OSSRS) procedure.

FIXED STEP SIZE RANDOM SEARCH (FSSRS)

The algorithm for FSSRS is

$$\boldsymbol{X}_{i+1} = \boldsymbol{X}_i - a_i \Delta \boldsymbol{X}_i + \Delta \boldsymbol{X}_{i+1} \tag{1}$$

where $\boldsymbol{X}_i$ is the position in state space at the ith instant and $\Delta \boldsymbol{X}_i$ is a random vector of length s, which is distributed uniformly over the hypersphere of radius s whose center is at the origin. The coefficient a_i is given by

$$a_i = \begin{cases} 0, & \text{if } Q_i < Q_{i-1}^+ \\ 1, & \text{if } Q_i \geq Q_{i-1}^+ \end{cases} \tag{2}$$

where

$$Q_i = Q(\boldsymbol{X}_i)$$

and

$$Q_{i-1}^+ = \min_{j=1,2,\cdots,i-1} Q_j$$

is the smallest value of the quality function to be observed in the first $i-1$ steps. The coefficient a_i serves to negate the effect of an unsuccessful step. Rastrigin[2],[3] has analyzed this algorithm for $s=1$ and the function

$$Q = \left[\sum_{i=1}^{n} x_i^2 \right]^{\frac{1}{2}}.$$

To motivate the development of a random search algorithm that adapts the step size s to the situation, the performance of FSSRS as a function of s is considered. Attention will be restricted to the function

$$Q(\boldsymbol{X}) = Q(\rho) = \sum_{i=1}^{n} x_i^2 = \rho^2 \tag{3}$$

which is smooth at its extremum and is more representative of problems with a minimum square error criterion than the function considered by Rastrigin. Otherwise, the analysis in this section is the same as Rastrigin.

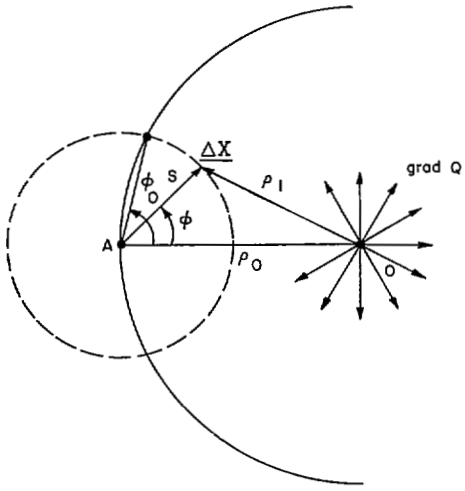

Fig. 1. A cross section of parameter space.

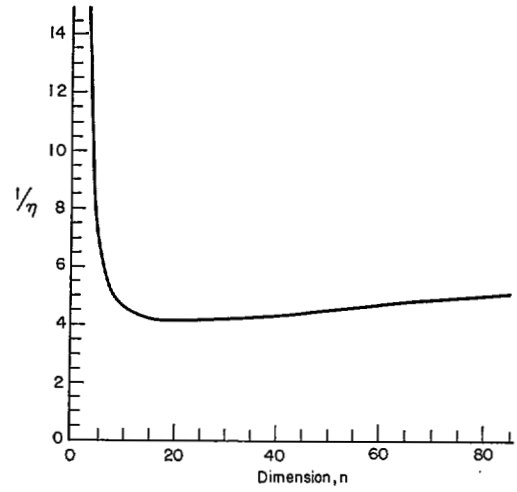

Fig. 2. Tradeoff between FSSRS and fixed step size gradient technique. The gradient method is superior above the boundary and FSSRS is superior below the boundary.

Consider the plane formed by the displacement vector ΔX and the gradient vector through the starting point A (Fig. 1). ϕ is the angle between the displacement vector and the negative gradient direction. ϕ_0 is the largest value for ϕ for which there is an improvement as the result of the step ΔX. For the assumed uniform distribution of displacement, the probability density for ϕ, considering ϕ only on $[0, \pi]$, (due to symmetry) is (see Rastrigin and Mutseniyeks,[4] and entries 858.45 and 858.46 of Dwight[5])

$$p(\phi) = \frac{\sin^{n-2}\phi}{2\int_0^{\frac{\pi}{2}} \sin^{n-2}\phi \, d\phi} = \frac{\Gamma(n-1)}{2^{n-2}\left[\Gamma\left(\frac{n-1}{2}\right)\right]^2} \sin^{n-2}\phi. \tag{4}$$

Search loss is defined as twice the ratio of the quality function to the expected value of the improvement per function evaluation. The search loss for the fixed step size gradient technique is then

$$L_g(n, \eta) = \frac{2(n+1)}{2\eta - \eta^2} \tag{5}$$

where $\eta = s/\rho$, the ratio of step size to distance to the minimum. The search loss for FSSRS is found to be

$$L_r(n, \eta) = \frac{2\int_0^{\frac{\pi}{2}} \sin^{n-2}\phi \, d\phi}{\eta\int_0^{\phi_0} \cos\phi \sin^{n-2}\phi \, d\phi - \frac{\eta^2}{2}\int_0^{\phi_0} \sin^{n-2}\phi \, d\phi} \tag{6}$$

and ϕ_0 is equal to $\cos^{-1}(\eta/2)$. Equations (5) and (6) are derived in Appendix I.

Fig. 2 shows the relative behavior of the random search and gradient methods for different values of η^{-1} and dimension n. Above the boundary random search is superior (has a smaller search loss) to the gradient method, while below the boundary the gradient technique is superior. For n less than 4, the gradient technique is always superior, but for higher dimension random search is superior for small η.

Optimum Step Size Random Search (OSSRS)

If the step size for FSSRS is very small, the probability of improvement is approximately one half, but the improvement is very small for a successful step, and this results in a small average improvement. On the other hand, if the step size is made too large, the step will overshoot the minimum and the probability of improvement will be extremely small, also resulting in a very small average improvement. Somewhere between these extremes lies an optimum step size, i.e., a step size for which the probability of the improvement of the quality function is not one half, but lies between zero and one half.

The expected improvement, normalized by the present value of Q, i.e., $I = -E\{\Delta Q\}/Q$, is equal to $2/L_r(n, \eta)$ and is given by

$$I(n, \eta) = \frac{\int_0^{\phi_0} (2\eta\cos\phi - \eta^2) \sin^{n-2}\phi \, d\phi}{2\int_0^{\frac{\pi}{2}} \sin^{n-2}\phi \, d\phi}. \tag{7}$$

To maximize I, the right-hand side of (7) is differentiated with respect to η and set equal to zero with the following equation for the optimal value of η resulting:

$$\int_0^{\phi_0} \sin^{n-2}\phi \, d\phi = \frac{(1 - \eta^2/4)^{\frac{n-1}{2}}}{n-1}, \quad 0 \le \eta \le 2. \tag{8}$$

Upon making the appropriate approximations, the following asymptotic expressions are found for large n (see Appendix II)

$$\eta_{\text{opt}} \sim \frac{1.225}{\sqrt{n}} \tag{9}$$

$$I_{\max} \sim \frac{0.406}{n} \tag{10}$$

$$p_{\text{opt}} \sim 0.27. \tag{11}$$

If the normalized expected improvement I is proportional to $1/n$ (see (10)), and if the normalized improvements are independent, as is true for OSSRS, then the average number of function evaluations for a fixed desired accuracy (relative to the starting value) is asymptotically linear in n, as is now demonstrated. Suppose that the search starts at a point at which the value of the quality function is Q_0, and that it is desired to terminate when the quality function reaches a final value of Q_f. The value Q_j of the quality function after j steps can be expressed recursively as

$$Q_j = Q_{j-1}(1 - i_j)$$

where i_j is the normalized improvement at the jth step. Thus

$$Q_M = Q_0 \prod_{j=1}^{M} (1 - i_j).$$

Taking the expected value of both sides, in light of the preceding assumptions, results in

$$E[Q_M] = Q_0\left(1 - \frac{k}{n}\right)^M \tag{12}$$

where k is the constant of proportionality. Solving for the value of M for which $E[Q_M]$ is equal to the desired final value Q_f results in

$$M = \frac{\log\left(\frac{Q_f}{Q_0}\right)}{\log\left(1 - \frac{k}{n}\right)}. \tag{13}$$

Thus the asymptotic expression for large n becomes

$$M \sim (\text{constant}) \cdot n \tag{14}$$

where the constant is equal to $(-1/k) \log (Q_f/Q_0)$.

Practical Algorithm for Adaptive Step Size Random Search (ASSRS)

OSSRS is a theoretical model and the optimum step size cannot be found without additional experimentation. One way to construct a practical algorithm would be to try numerous exploratory random steps from the same point, each with the same step size, and to repeat this procedure for a number of different step sizes. From these results, the optimum step size could be estimated and this estimate used. However, none of the intermediate exploratory steps would produce any improvement. In the ASSRS algorithm, no attempt is made to estimate the optimum step size accurately. Instead the optimum is tracked in an approximate fashion, and each step is both exploratory and able to produce an improvement. A nominal value, s, for the step size is chosen before each iteration. A random step of size s is taken and a random step of size $s(1+a)$ is taken $(1>a>0)$, and the resultant normalized improvements are compared. The step size that produces the larger improvement is chosen as the nominal step size for the next iteration. If neither step causes an improvement, the step size remains unchanged; and if this occurs for some number of iterations, the step size is reduced. Thus on the average the algorithm adjusts to the direction of the best step size. In addition, each time some large number of iterations has passed, a step with nominal step size is compared, in the same manner, with a step of much larger size. Again, the step size that produces the larger improvement is chosen as the new nominal step size. This test serves as a deterrent against the possibility that the step size has inadvertently become too small. It is also helpful in the case in which the I vs. step size curve has more than a single local maximum. In such a case the search procedure could be chasing a small local maximum; and a large change in step size would make it possible to detect and begin adaptation to a higher local maximum. A flow diagram for the ASSRS algorithm is shown in Fig. 3.

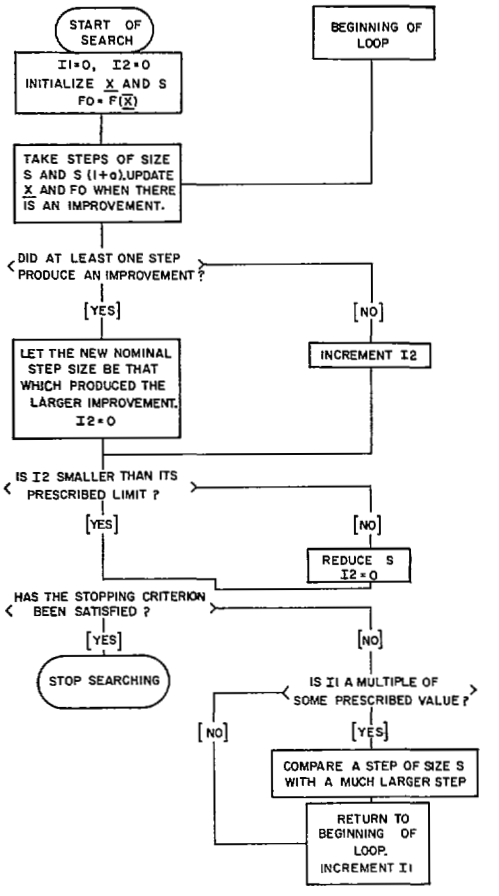

Fig. 3. Flow diagram for ASSRS. I2 counts the number of successive failures. I1 counts total number of iterations through the loop.

Experimental Results

The ASSRS algorithm was tested on the IBM 7094 computer for a number of test functions. The results of these experiments are presented here.

The quality function $Q=\rho^2$ was tested to see how the ASSRS algorithm compared to the Newton–Raphson method and to OSSRS. For each dimension from 1 to 40, fifteen independent trials were run. The stopping criterion was $Q<10^{-8}$ and the starting point was $(1, 1, \cdots, 1)$. The resulting average number of required function evaluations is well described by $\text{ICALL}=80n$.[1]

Since derivatives are not available, partial derivatives must be measured approximately by taking finite differences, and the number of function evaluations per iteration for the Newton–Raphson method can be found as follows. To approximate the gradient vector, a minimum of n finite differences are needed. To find the diagonal terms of the Hessian matrix, $2n$ additional function evaluations are needed to estimate the partial derivatives at a second point. The Hessian matrix is symmetric so that $h_{ij}=h_{ji}$. Thus one half of the off-diagonal terms or $(n^2-n)/2$ more second partial derivatives are required. This requires n^2-n additional function evaluations. Also, one more function evaluation occurs with the final move. Thus the total number of function evaluations required per iteration of the Newton–Raphson method is $n+2n+(n^2-n)+1=(n+1)^2$. Assuming that the partial derivatives can be determined exactly by finite differences, only one iteration is required for this function and the number of function evaluations required for the Newton–Raphson method is $(n+1)^2$.

The results for ASSRS and for the Newton–Raphson method are shown in Fig. 4. Extrapolating these curves, their intersection is at $n=78$, beyond which dimension ASSRS is superior. If the fact that the partial derivatives cannot be measured exactly (which means that additional iterations are required) is taken into account, the Newton–Raphson curve becomes higher and the intersection with the ASSRS occurs at a smaller value of n. Also, if a lesser degree of accuracy is required, the ASSRS curve has a smaller slope while the Newton–Raphson curve is unchanged. This also results in a smaller value of n at the intersection of the curves, and ASSRS is then superior at a smaller value of dimension. As an indication of the variance in ICALL associated with ASSRS, the standard deviation for $n=100$ was found to be 502 as compared with the mean of 7677.

On the basis of the experimental results, the average normalized improvement per step, I, was calculated for $Q=\rho^2$ and was found to be asymptotic to $0.2725/n$. Thus, although the value of k is smaller than for OSSRS, the asymptotic form of I is still k/n; and the number of function calls was found to be asymptotically

[1] The variable ICALL represents the number of function evaluations performed during a given minimization procedure.

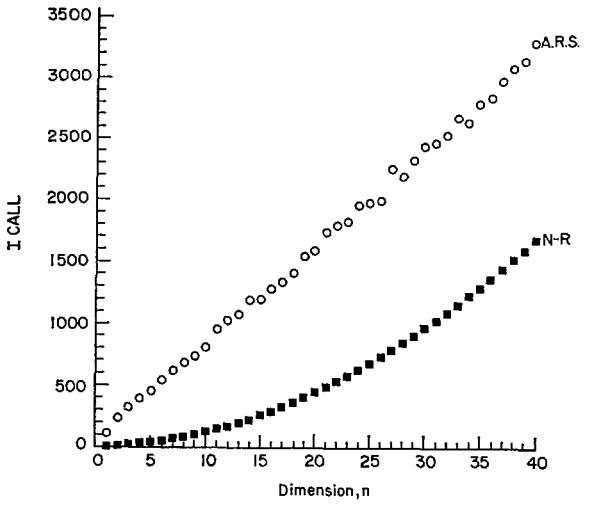

Fig. 4. Average number of function evaluations vs. dimension for $Q=\sum_{i=1}^{n}x_i^2$, for ASSRS (O) and the Newton–Raphson method (squares).

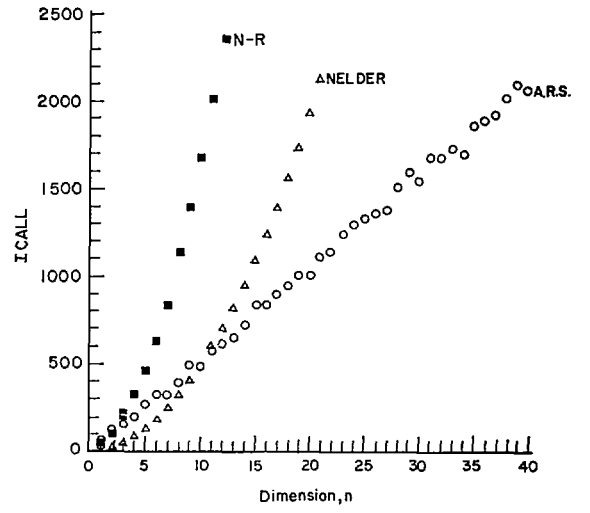

Fig. 5. Average number of function evaluations vs. dimension for $Q=\sum_{i=1}^{n}x_i^4$, for ASSRS (O), Newton–Raphson (squares), and the simplex method (Δ).

proportional to n as it is for OSSRS [see (14)].

Fig. 5 shows the results for

$$Q=\sum_{i=1}^{n}x_i^4$$

averaged for 22 independent experiments for each dimension from 1 through 40. Again the starting point was $(1, 1, \cdots, 1)$. The stopping criterion was $Q<0.5\times10^{-8}$. The Newton–Raphson method is again assumed to be able to measure derivatives without error, but the fact that Q is not quadratic is taken into account by multiplying the number of iterations required by $(n+1)^2$ in order to find the number of function calls. The results for the simplex method are those of Nelder and Mead,[6] who found that, for $n=1$ through 10, the number of function evaluations needed for the simplex method is well described by $3.16(n+1)^{2.11}$. This formula was extrapolated to higher dimensions and

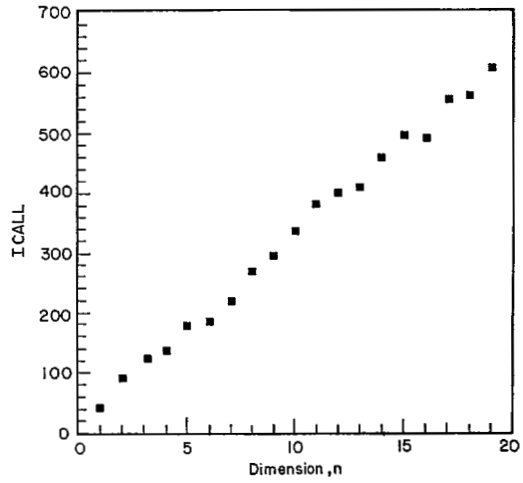

Fig. 6. Average number of function evaluations vs. dimension for ASSRS for $Q=\sum_{i=1}^{n} a_i x_i^2$.

plotted along with the curves for ASSRS and Newton–Raphson. In this case, ASSRS is superior to Newton–Raphson for $n>2$ and to the simplex method for $n>10$. For ASSRS, ICALL is again proportional to n, and I is asymptotic to $0.427/n$.

The quadratic form

$$Q = \sum_{i=1}^{n} a_i x_i^2$$

was also used as a test function, with the coefficients $a_1, a_2, \cdots, a_n$ randomly chosen from a uniform distribution on $[0.1, 1]$. The minimization was repeated twenty times for each dimension from 1 to 20 and six times for $n=100$. The search was terminated when Q was less than one one-thousandth of its initial value. Fig. 6 shows the resulting average number of function evaluations required as a function of dimension. The result for 100 dimensions was ICALL = 3396. Again, ICALL was found to be approximately linear in n.

Rastrigin[3] has done some work with random search algorithms that adapt only to the best direction with a fixed step size. The problem with this method is that convergence is only guaranteed to within a distance of one step size to the minimum, so that if a high accuracy is desired, the steps must be small, thus requiring many function evaluations. Adapting the step size seems to be more fruitful than adapting to the direction. A combination of the two methods, however, would seem to be in order and should be the subject of future investigations.

ASSRS was also tested using Rosenbrock's function $Q=100(x_2-x_1^2)^2+(1-x_1)^2$, and although it converged, ASSRS was inferior to Rosenbrock's method[7] and to Powell's method[8] for this function. It should, therefore, be noted that ASSRS is not very effective as a ridge follower, but shows its superiority in multidimensional problems without narrow valleys or ridges. Combining directional adaptation with step size adaptation may result in removing this limitation.

Conclusion

In this paper the problem of minimizing a function of several parameters by the method of random search has been discussed. The fixed step size random search algorithm has been described and compared to the simple gradient technique on the basis of search loss. It has been shown that for $n>4$, FSSRS is superior for sufficiently small η. Optimum step size random search was introduced and its performance investigated for hyperspherical surfaces.

A practical algorithm for adaptive step size random search has been described and compared with deterministic methods. For the functions

$$Q = \sum_{i=1}^{n} x_i^2$$

$$Q = \sum_{i=1}^{n} x_i^4$$

$$Q = \sum_{i=1}^{n} a_i x_i^2$$

the number of function evaluations required for a desired accuracy for the deterministic methods increases at a rate that is proportional to at least the second power of n; and the computation time increases as n^3. This is true for other classical methods besides the Newton–Raphson and simplex methods.[9] However, for ASSRS the number of required function evaluations is proportional to n, and the computation time is proportional to n^2. The computation time could be made proportional to n if parallel computations were used. Thus the conclusion is reached, that despite its simplicity, adaptive random search is an attractive technique for problems with large numbers of dimensions.

Appendix I

Evaluation of Search Loss

The search loss for the fixed step size gradient method is found as follows. The gradient of Q is given by $2\boldsymbol{X}$; thus the step of size s to be taken is

$$\Delta \boldsymbol{X} = \frac{-s \operatorname{grad} Q}{|\operatorname{grad} Q|} = \frac{-s}{\rho} \boldsymbol{X}.$$

This results in a change in Q

$$\Delta Q = s^2 - 2s\rho \tag{15}$$

which is negative as long as $\rho>s/2$. To determine the correct descent direction, n measurements of the quality function are made, i.e., one in each of the n coordinate directions with a step size much smaller than s. Also, one function evaluation is made corresponding to the actual move of size s. Thus a total of $n+1$ function evaluations are necessary for the improvement given by (15). Therefore, the search loss for the gradient method is

$$L_g(n, \rho) = \frac{2\rho^2(n+1)}{2s\rho - s^2} \cdot \tag{16}$$

Subsequent substitution of (19), (21), and (22) into (18) yields the final expression for the search loss for the random search method:

$$L_r(n, \rho) = \frac{2\int_0^{\frac{\pi}{2}} \sin^{n-2}\phi\, d\phi}{\frac{s}{\rho}\int_0^{\phi_0(\rho)} \cos\phi \sin^{n-2}\phi\, d\phi - \frac{s^2}{2\rho^2}\int_0^{\phi_0(\rho)} \sin^{n-2}\phi\, d\phi} \cdot \tag{23}$$

For the random search method, reference is again made to Fig. 1. Because of the symmetry of the quality function, only positive values of ϕ need to be considered in this discussion. Let ϕ_0 be the angle subtended by all possible successful steps of size s and let ρ_0 be the initial distance to the minimum. Then the relation between ϕ_0 and ρ_0 is

$$\phi_0 = \cos^{-1}\frac{s}{2\rho_0} \cdot \tag{17}$$

The search loss for FSSRS is

$$L_r(n, \rho) = \frac{2\rho^2}{p\langle \Delta Q\rangle} = \frac{2\rho^2}{p\langle(\Delta\rho^2)\rangle} \tag{18}$$

where p is the probability of a successful step. $\langle\cdot\rangle$ denotes mathematical expectation over all successes, and $\Delta\rho^2$ is the difference between the initial and final distances to the minimum for one step of the random search. Because steps that are unsuccessful are rejected, $\Delta\rho^2 = 0$, and there is no improvement. Referring to Fig. 1,

$$\begin{aligned}\Delta\rho^2 &= -\Delta Q = \rho_0^2 - \rho_1^2 \\ &= \begin{cases} 2s\rho\cos\phi - s^2, & \text{if } |\phi| < \phi_0 \\ 0, & \text{if } |\phi| \geq \phi_0 \end{cases}\end{aligned} \tag{19}$$

where $\rho = \rho_0$. The expectation of $\Delta\rho^2$ over all successful steps is

$$\langle\Delta\rho^2\rangle = \frac{\int_0^{\phi_0(\rho)} \Delta\rho^2 p(\phi)\, d\phi}{p} \cdot \tag{20}$$

The probability of success p is, from (4), given by

$$\begin{aligned} p &= \text{probability of success} \\ &= \frac{\int_0^{\phi_0(\rho)} \sin^{n-2}\phi\, d\phi}{2\int_0^{\frac{\pi}{2}} \sin^{n-2}\phi\, d\phi} \cdot \end{aligned} \tag{21}$$

Substitution of (4) and (21) into (20) gives

$$\langle\Delta\rho^2\rangle = \frac{\int_0^{\phi_0(\rho)} \Delta\rho^2 \sin^{n-2}\phi\, d\phi}{\int_0^{\phi_0(\rho)} \sin^{n-2}\phi\, d\phi} \cdot \tag{22}$$

Letting $\eta = s/\rho$, (16) and (23) become (5) and (6), respectively, and ϕ_0 is now given by substituting into (17):

$$\phi_0 = \cos^{-1}\frac{\eta}{2} \cdot$$

Appendix II

Asymptotic Expressions for η_{opt}, I_{max}, and p_{opt}

The asymptotic expressions for large n will now be derived. For simplicity, in the subsequent calculations, the optimum values for p, η, and I will simply be written as p, η_0, and I, respectively. The first assumption that will be made is that for large n, η_0 is small, which assumption is borne out by the results obtained by evaluation of (7). It is, therefore, assumed that for large n, $\sin^{-1}\eta_0/2 \approx \eta_0/2$. If the expression

$$\int_0^{\frac{\pi}{2}} \sin^{n-2}\phi\, d\phi$$

is denoted by $a(n)$, (8) becomes

$$\eta_0 a(n) - \int_{\frac{\pi}{2}-\frac{\eta_0}{2}}^{\frac{\pi}{2}} \sin^{n-2}\phi\, d\phi \approx \frac{\left(1 - \frac{\eta_0^2}{4}\right)^{\frac{n-1}{2}}}{n-1} \cdot \tag{24}$$

Using the identity

$$\int_0^{\frac{\eta_0}{2}} \cos^n\phi\, d\phi = \int_{\frac{\pi}{2}-\frac{\eta_0}{2}}^{\frac{\pi}{2}} \sin^n\phi\, d\phi$$

(24) becomes

$$\eta_0 a(n) - \int_0^{\frac{\eta_0}{2}} \cos^{n-2}\phi\, d\phi \approx \frac{\left(1 - \frac{\eta_0^2}{4}\right)^{\frac{n-1}{2}}}{n-1} \cdot \tag{25}$$

η_0 was assumed to be small. Therefore, all values of ϕ in the integral on the left-hand side of (25) are small, and the approximation $\cos\phi \approx 1 - \phi^2/2$ is assumed to be valid. Also, because η_0 is small the expression on the right-hand side of (25) is approximated by the first three nonzero terms of its Taylor series expansion, viz.,

$$\frac{1}{n-1} - \frac{\eta_0^2}{8} + \frac{(n-3)\eta_0^4}{128}.$$

Upon making these substitutions (25) becomes

$$\eta_0 a(n) - \eta_0\left[\frac{\eta_0}{2} - \frac{\eta_0^3(n-2)}{48}\right] \approx \frac{1}{n-1} - \frac{\eta_0^2}{8} + \frac{(n-3)\eta_0^4}{128}. \tag{26}$$

Using Dwight's entry 858.44[5] and Stirling's approximation for $n!$, the asymptotic expression for $a(n)$ becomes

$$a(n) \approx \sqrt{\frac{\pi}{2n}}. \tag{27}$$

After substituting $\sqrt{\pi/2n}$ for $a(n)$ and assuming that n is large, (26) becomes

$$\frac{5}{384}\eta_0^4 n - \frac{3}{8}\eta_0^2 + \sqrt{\frac{\pi}{2n}}\eta_0 - \frac{1}{n} \approx 0. \tag{28}$$

Solving (28) for η_0, the resultant asymptotic expression for η_0 is, as given in (9),

$$\eta_{opt} \approx \frac{1.225}{\sqrt{n}}.$$

Noting that

$$\int_0^{\cos^{-1}\frac{\eta_0}{2}} \sin^{n-2}\phi\, d\phi = 2a(n)p$$

(23) can be rewritten as

$$2\eta_0 a(n)p \approx \frac{\left(1 - \frac{\eta_0^2}{4}\right)^{\frac{n-1}{2}}}{n-1}.$$

Solving for p,

$$p \approx \frac{\left(1 - \frac{0.375}{n}\right)^{\frac{n-1}{2}} \cdot n}{3.07(n-1)}.$$

The asymptotic expression for large n is, therefore,

$$p_{opt} \approx 0.27$$

as in (11). Recalling (7) and (8), the expression for optimum I becomes

$$I = \frac{\eta_0^2 \int_0^{\cos^{-1}\frac{\eta_0}{2}} \sin^{n-2}\phi\, d\phi}{2a(n)} = \eta_0^2 p.$$

Upon substituting the asymptotic expressions previously obtained for p and η_0 the result is (10), viz.,

$$I_{max} = \frac{0.406}{n}.$$

References

[1] S. H. Brooks, "Discussion of random methods for locating surface maxima," *Operations Research*, vol. 6, pp. 244–251, 1958.

[2] L. A. Rastrigin, "The convergence of the random search method in the extremal control of a many parameter system," *Automation and Remote Control*, vol. 24, pp. 1337–1342, 1963.

[3] —— *Sluchainyi Poisk v Zadachakh Optimizatsii Mnogoparametricheskikh Sistem*. Riga, USSR: Akademiia Nauk Latviiskoi SSR, 1965.

[4] L. A. Rastrigin and V. A. Mutseniyeks, "Extremal control of continuous multiparameter systems by the method of random search," *Engrg. Cybernetics*, vol. 2, pp. 82–90, January/February 1964.

[5] H. B. Dwight, *Tables of Integrals and Other Mathematical Data*, 4th ed. New York: Macmillan, 1961.

[6] J. A. Nelder and R. Mead, "A simplex method for function minimization," *Computer J.*, vol. 7, pp. 308–312, January 1965.

[7] H. H. Rosenbrock, "An automatic method for finding the greatest or least value of a function," *Computer J.*, vol. 3, pp. 175–184, October 1960.

[8] M. J. D. Powell, "An iterative method for finding the stationary values of a function of several variables," *Computer J.*, vol. 5, pp. 147–151, July 1962.

[9] D. J. Wilde, *Optimum Seeking Methods*. Englewood Cliffs, N. J.: Prentice-Hall, 1964.

Michael A. Schumer (S'62–M'68) was born in Brooklyn, N. Y., on March 27, 1943. He received the B.E.E. degree from the City College of the City University of New York, New York, N. Y., in 1964, and the M.A. and Ph.D. degrees in electrical engineering from Princeton University, Princeton, N. J., in 1966 and 1967, respectively.

From 1964 to 1965, he was a Member of the Technical Staff of Bell Telephone Laboratories, Inc., Holmdel, N. J. From 1965 to 1967, he completed his graduate studies at Princeton University as a National Science Foundation Graduate Fellow. Since November, 1967, he has been a Senior Research Scientist in the Research Div., Raytheon Co., Waltham, Mass., where he is concerned with basic and applied research in communication theory.

Dr. Schumer is a member of Eta Kappa Nu.

Kenneth Steiglitz (S'57–M'64) was born in Weehawken, N. J., on January 30, 1939. He received the B.E.E. degree in 1959, the M.E.E. degree in 1960, and the Eng.Sc.D. degree in 1963, all from New York University, New York, N. Y.

In 1963 he joined the Dept. of Electrical Engineering, Princeton University, Princeton, N. J., where he is now Associate Professor.

Dr. Steiglitz is a member of Eta Kappa Nu, Tau Beta Pi, and Sigma Xi.